



Multimodal data-driven machine learning for the prediction of surface topography in end milling

L. Hu^{1,2} · H. Phan³ · S. Srinivasan³ · C. Cooper⁴ · J. Zhang⁴ · B. Yuan³ · R. Gao⁴ · Y. B. Guo^{1,2} 

Received: 6 July 2023 / Accepted: 7 December 2023 / Published online: 9 January 2024
© The Author(s) under exclusive licence to German Academic Society for Production Engineering (WGP) 2024

Abstract

Prediction of surface topography in milling usually requires complex kinematics and dynamics modeling of the milling process, plus solving physical models of surface generation is a daunting task. This paper presents a multimodal data-driven machine learning (ML) method to predict milled surface topography. The proposed method predicts the height map of the surface topography by fusing process parameters and in-process acoustic information as model inputs. This method has been validated by comparing the predicted surface topography with the measured data.

Keywords Machine learning · Multimodal data · Surface topography · Milling

1 Introduction

1.1 Background

End milling plays a key role in manufacturing functional components in industries. Compared with alternative processes such as electrical discharge machining (EDM) and grinding, end milling can create a variety of complex geometries with high efficiency. Milling also generates a relatively smooth surface and thus can be used as a post-process for components made by metal additive manufacturing [1]. With these advantages, milling has become a competitive manufacturing process in fabricating turbine blades, dies, and molds [2]. Surface finish by the milling process can be affected by many factors, e.g., milling parameters [3], cutting fluids [4, 5], cutting tool conditions [6], and vibrations [7].

As a material removal process, milling generates surface topography [8], which directly affects surface functionality, such as friction, wear, and fatigue [9]. Topographical roughness measures, such as arithmetical mean height S_a and root mean square height S_q , are critical when surface functionality is concerned. For instance, surface topography controls the tribology, corrosion, and fatigue of mechanical components [10, 11]. Additionally, surface topography controls mass transport, cell adhesion, and bone ingrowth of orthopedic implants, which determine the effectiveness and useful life of these crucial medical devices [12]. Therefore, surface topography prediction in the milling process is necessary to improve surface quality and efficiency. Assessing the surface topography of each component in real-world scenarios can be time-consuming and expensive. For instance, the automotive industry manufactures hundreds of thousands of vehicles every month [13]. In light of such circumstances, manually evaluating surface topography for each component is impractical. Consequently, there is a practical need to develop a cost-effective and efficient method to predict milled surface topography rapidly.

1.2 Physical methods to predict surface topography

Milled surface topography depends on various factors such as material properties, cutting force and vibrations [14], cutting tool geometry [15], and coolant [5]. Current methods for surface topography prediction can be roughly categorized into physical methods and data-driven methods.

✉ Y. B. Guo
yuebin.guo@rutgers.edu

¹ Department of Mechanical and Aerospace Engineering, Rutgers University-New Brunswick, Piscataway, NJ 08854, USA

² New Jersey Advanced Manufacturing Institute, Rutgers University-New Brunswick, Piscataway, NJ 08854, USA

³ Department of Electrical and Computer Engineering, Rutgers University-New Brunswick, Piscataway, NJ 08854, USA

⁴ Department of Mechanical and Aerospace Engineering, Case Western Reserve University, Cleveland, OH 44106, USA

Physical methods predict surface topography using physical laws underlying the milling process. The physical models typically use initial milling process parameters and conditions to predict process phenomena such as chatter, tool wear, surface roughness, etc. Physical methods can be divided into two subcategories: analytical and numerical methods. Table 1 summarizes the physical methods of milling analysis in the literature.

Montgomery and Altintas [14] described a method to predict milling forces and surface topography through kinematics modeling of milling tools and workpiece vibrations. Altıntaş and Budak [16] proposed a method using transfer functions of the milling tool in a stability model to predict milling stability analytically. Schmitz et al. [17] considered the runout effect of the milling tool on the workpiece surface and predicted the surface line profile based on a milling force model. Both Altintas et al. and Ding et al. [7, 19] proposed similar discrete time-domain methods to predict the milling process stability lobes and analyze the surface finish. Arizmendi et al. [18] proposed a surface topography and roughness prediction method by applying tool vibration on a geometrical surface generation model. Altintas and Jin [20] predicted the milling forces using a slip-line field model considering material properties and friction coefficient. Kaymakci et al. [22] proposed a model to predict machining forces using the cutting insert geometry and milling parameters. The model is then extended to boring, turning, and other machining operations using cutting inserts. Costes and Moreau [21] showed an analytical model for 3D surface topography prediction considering the tool vibration.

The prediction of tool wear rate and milling forces is prevalent in numerical methods. Rao et al. [23] used the finite element method (FEM) with a tool-workpiece contact 3D model in the milling process to analyze tool wear

patterns. This method also investigates the influence of tool wear on surface integrity. Jin and Altintas [24] used a tool-workpiece contact 3D model in FEM to predict milling forces based on cutting tool geometry, material properties, and initial milling conditions. Wu and Zhang [25] presented a tool contact model considering material damage and heat transfer to predict milling forces. Felhő et al. [26] presented a FEM approach with cutting insert geometry and initial milling parameters to predict surface topography.

The current physical methods have been widely used to predict critical milling parameters such as milling forces, process stability, tool wear, and surface finish. However, the physical methods require the milling process to be mathematically modeled, which could be challenging and demanding to implement in real-life production. Not to mention, building the mathematical model itself could be challenging and sometimes hard to validate. Physical methods could be used to generate data for analysis or prediction to reduce the number of physical experiments, but it is also hard to verify the synthesized data quality and the accuracy of analysis or prediction results against actual manufacturing environments. Given these challenges, data-driven methods have been increasingly explored for topography prediction.

1.3 Data-driven methods to predict surface topography

Data-driven methods refer to approaches that use pure data to train machine learning (ML) models for prediction. Such methods usually involve massive datasets for training purposes. Table 2 summarizes some recent data-driven methods for machining processes in the literature.

Yang and Liu [29] used multiple linear regression to predict the influence of different milling parameters on surface

Table 1 Physical models for milling analysis

Model type	Model inputs	Model outputs	Refs.
Analytical	Milling parameters, tool geometry	Milling forces, surface topography	[14]
	Milling parameters, tool geometry (contact zone), system natural frequency	Milling stability	[16]
	Milling parameters, milling forces, system natural frequency	Line profile	[17]
	Milling parameters, system natural frequency	Milling stability, milling forces	[7]
	Milling parameters, tool geometry, tool vibration	Surface topography	[18]
	Milling parameters, system natural frequency	Milling stability	[19]
	Milling parameters, material properties	Milling forces	[20]
	Milling parameters, tool geometry, tool displacement	Line profile	[21]
	Milling parameters, tool geometry, cutting angles	Milling forces	[22]
Numerical	Milling parameters, material properties, tool geometry	Surface integrity, specific cutting energy, tool wear	[23]
	Milling parameters, material properties, tool geometry, tool-workpiece interaction	Milling forces	[24]
	Milling parameters, material properties, tool geometry	Milling forces	[25]
	Milling parameters, tool geometry	Surface topography	[26]

Table 2 Data-driven methods for milling surface prediction

Model type	Model inputs	Model outputs	Refs.
Genetic algorithms	Milling parameters, tool geometry	Surface roughness	[27]
Genetic algorithm	Milling parameters	Surface roughness	[28]
Linear regression	Milling parameters	Surface roughness	[29]
Bayesian inference	Milling parameters	Milling stability	[30, 31]
FCNN	Milling parameters	Surface roughness	[32]
SVM	Cutting forces	Surface topography, milling stability	[33]
SVM, ELM	EDM process parameters	Surface roughness	[34]
CNN	Surface topography	Surface roughness	[35]
GAN	Turning parameters, tool vibration, turning forces	Surface image	[36]
CNN	Milling vibrations	Milling stability (chatter)	[37]
Classification tree	Milling acoustic signals	Surface roughness	[38]
CNN	Milling acoustic signals	Part defect (pores)	[39]
GAN, LSTM	Milling vibration and acoustic signals	Tool wear	[40]
CNN	Milling acoustic signals	Tool wear	[41]

roughness. Suresh Kumar Reddy and Venkateswara Rao Suresh [27] applied genetic algorithms for milling roughness prediction, where the input is milling parameters with 81 experiments with four 2-level parameters, and the output is surface roughness. Karandikar et al. [30] employed a Bayesian-based learning algorithm and milling parameters (data size of 105 samples) to analyze milling stability. Çolak et al. [28] utilized 84 datasets and genetic expression programming for surface roughness prediction, where the input is milling parameters, and the output is surface roughness. Zain et al. [32] adopted a fully connected neural network (FCNN) and 24 datasets to predict surface roughness. Liu et al. [35] presented a convolutional neural network (CNN) to estimate surface roughness parameters based on surface topography images. Paturi et al. [34] used 81 experiments on four different models to predict the surface roughness of the EDMed workpiece. Peng et al. [33] developed a support vector machine (SVM) model using milling parameters and vibration signals during the cutting process to generate the milled surface topography and identify stable cutting process windows. Rahimi et al. [37] presented a CNN-based model which used milling vibration signals to predict the process stability.

As the acoustic signal is an easy-to-acquire process signature indicative of surface generation, it is widely used in data-driven approaches to predict tool and workpiece contact conditions. The acoustic signal originates from the varying tool-chip interface, and the changes to the interface are reflected as changes in the signal. Gauder et al. [39] presented a CNN-based deep-learning model using acoustic signals for the prediction of workpiece pores in milling. Shah et al. [40] proposed a method using LSTM and GAN to analyze milling vibration and acoustic emission signals for the prediction of tool wear. Griffin et al. [38] proposed a method

using acoustic data and short-time Fourier transform (STFT) to predict surface parameters. Cooper et al. [41] presented a CNN model for the classification of tool conditions using acoustic signals. Although the acoustic signal is popular in analyzing tool wear and workpiece defects, few data-driven methods use acoustic signals for topography prediction.

Cao et al. [36] proposed a surface topography generation method in a turning process using a generative adversarial network (GAN). The inputs of the GAN are the turning parameters, tool vibration signal, and turning forces. The outputs are grayscale images of turning surfaces. However, the generated 2D images can only be used for visualization as they contain no height data. Moreover, this method uses the conventional SRResNet architecture [42], which is less efficient and produces low-resolution images compared to other popular GAN architectures. In summary, although acoustic signals are often used in analyzing tool wear and workpiece quality, there are few data-driven methods using acoustic signals for topography prediction.

1.4 Hybrid physics-informed data-driven approaches

It is essential to acknowledge the emerging hybrid methods, which combine the advantages of both the physics-driven and data-driven approaches to enhance model performance. One example of the hybrid method is Physics-Informed Neural Networks (PINNs), which integrate neural networks with physical laws to solve partial differential equations (PDEs). Such a method is very important in solving heat transfer [43] and fluid mechanics [44] problems. PINNs leverage the strengths of both data-driven and physics-based methodologies to enable more accurate predictions. Despite the promising potential, the application of PINNs in the manufacturing

domain, e.g., machining, is still in its infancy [45]. Further research and development are needed to explore the use of PINNs to predict surface topography.

1.5 Research objectives

The literature analysis has shown that, on the one hand, physical methods may predict milling surface finish by leveraging machining process conditions such as tool wear, process stability, and cutting forces. However, the current physical methods require complex mathematical modeling of the milling process, which is often very challenging to derive and implement. On the other hand, data-driven methods can predict tool wear, milling stability, and surface attributes (e.g., roughness) based on the machining process conditions. The data-driven approach can also be a key component of digital twins [46]. Nevertheless, the data-driven approaches for predicting milled surface topography are still at an early stage. Some methods can predict surface images for visualization, but few can accurately predict surface topography.

To fill the knowledge gap and achieve more comprehensive and accurate results in predicting milling surface topography, an in-depth multimodal ML method has been developed in this work. The characteristics of the developed method are summarized as follows:

- The proposed approach predicts surface topography using an ML method. Current data-driven methods for roughness quantification focus on scalar roughness prediction, e.g., S_a , or generate surface images. However, these scalars and images cannot fully model the intricate topography induced by milling.
- The proposed method uses multimodal data by fusing the machining process conditions and milling parameters. The bulk of ML models is limited to a single type of data modality for surface analysis.

- The proposed method is based on the Generative Adversarial Network (GAN) [47]. GAN is a widely used image generation framework that can effectively learn the statistics and characteristics of the input data and generate accurate output based on the conditioned input.

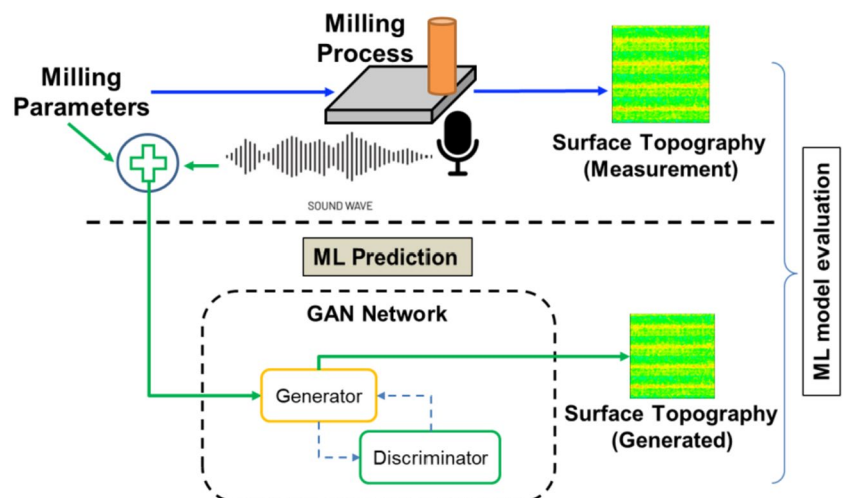
The paper is structured as follows: Sect. 2 presents the research route, milling experiments, and data pre-processing. Section 3 proposes a GAN-based ML model for surface generation. Section 4 presents and discusses the ML output by comparing model outputs with experimentally measured surfaces. Section 5 summarizes the key results of this work and presents future outlooks.

2 Milling experiment and data ingestion

2.1 Research route

The overall research route is shown in Fig. 1. The milling parameters, i.e., cutting speed (v), feed/tooth (f_z), radial depth-of-cut (DoC) a_e , and axial DoC a_p , are first planned based on the Design of Experiment (DoE) method and then programmed into G-codes to conduct the milling experiments. A microphone was placed next to the workpiece inside the milling chamber to record the milling acoustic signals. After milling, the machined surfaces are measured using the Keyence 3D measurement system VR-3000 to obtain surface topography. On the other hand, the milling parameters and the recorded acoustic data will be used as multimodal inputs to train the GAN model. The output of the GAN is surface topography generated by the machine learning (ML) model. The generated surface topography and measured surface topography will be compared to conduct a post-analysis to study the accuracy and cost of the ML model.

Fig. 1 Research route



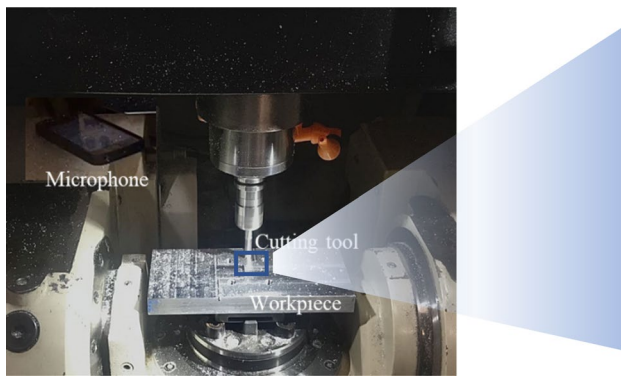


Fig. 2 Milling experiment setup

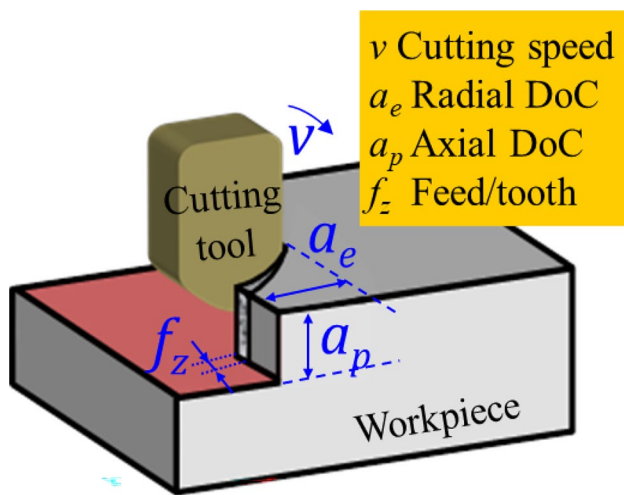


Fig. 3 Material removal schematics in milling

2.2 Milling experiment design

The experiment setup is shown in Fig. 2 with the close-up view of the tool/workpiece interaction zone in Fig. 3. The milling experiments were performed on the Mircolution 5100 High-Speed 5-Axis CNC machine with a 2-flute uncoated carbide 6.35 mm flat end milling tool. The work material is Aluminum alloy 6061-T6. Acoustic signals were recorded in real-time with a microphone sampling at 48 kHz during milling. The microphone is placed next (~300 mm) to the workpiece at the same location for each milling experiment. The machined surface topographies were measured using the Keyence VR-3000 measurement system.

The milling process involves several stages: air cutting, tool engagement, stable cutting, cutting direction change, and tool disengagement. Figure 4 shows a schematic tool path of the milling process, where the yellow line is the milling tool path, and the gray area is the workpiece. At the beginning of the milling process, the tool path starts

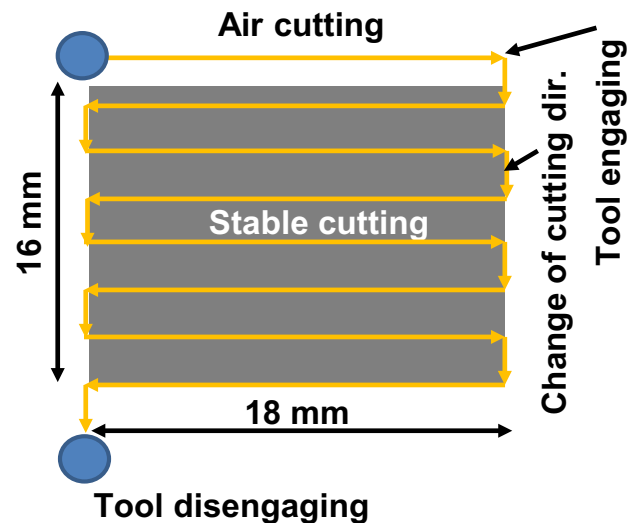


Fig. 4 Milling process stages

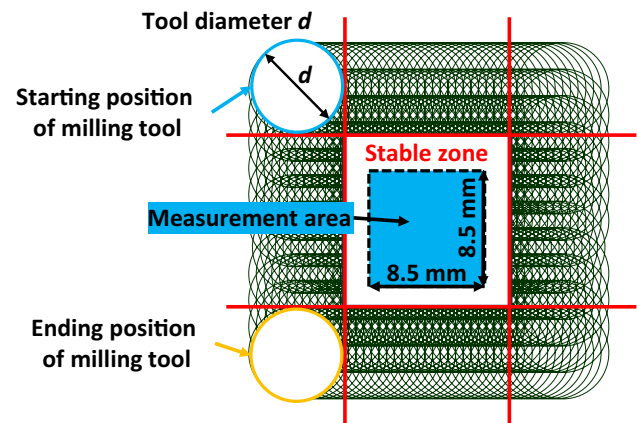


Fig. 5 Transitional and stable milling zones

at the top-left position, and the tool does not contact the workpiece. Each milled surface is a 16 mm × 18 mm rectangle. When the milling process starts, the rotating tool moves along the yellow line without contacting the workpiece (i.e., air cutting). On the right side, the path arrow pointing to the workpiece represents the tool's engagement with the workpiece. Stable cutting starts when the milling tool is fully engaged with the workpiece. The milling tool changes its cutting direction at the workpiece edges. When the milling process is finished, the milling tool disengages the workpiece. Figure 5 [48] shows the stable cutting zone, situated at a distance of one tool diameter (d) from the milled surface edges. Surface measurements were conducted within the stable zone of 8.5 mm × 8.5 mm.

To cover the wide range of cutting conditions, the milling parameters are planned in the ranges as shown in Table 3a.

Table 3 Milling parameters and ranges

(a) ISO milling parameters			(b) G-code parameters		
4-Parameter	Range of 6-level	Increment	4-Parameter	Range of 6-level	Increment
Cutting speed	120–320 m/min	40 m/min	Spindle speed	6000–16,000 rpm	2000 rpm
Feed/tooth	0.025–0.035 mm	0.002 mm	Feed rate	300–1100 mm/min	–
Axial DoC	0.5–1.5 mm	0.2 mm	Axial DoC	0.5–1.5 mm	0.2 mm
Radial DoC	0.2–1.2 mm	0.2 mm	Radial DoC	0.2–1.2 mm	0.2 mm

The first column shows the milling parameters, and the second column indicates the ranges of the parameters. The increment shows the difference between the two adjacent levels. For example, the cutting speed starts at 120 m/min and then increases by 40 m/min, so the next level cutting speed is 160 m/min. To cover more possible combinations, we chose the four parameters in six levels, which resulted in 6^4 (i.e., 1296) different parameter combinations. Each process parameter value has been used 216 times in the parameter space, which would avoid the need to repeat the experiment and the associated high costs. Other setups, such as 5-level or 7-level experimental designs, produce either inefficient numbers of samples (< 650) or would be too expensive to produce. Note that cutting speeds and feeds/tooth are converted into spindle speeds and feed rates in the G-code, respectively. It is worth mentioning that the tool condition was examined throughout the experiments, and tool wear was measured using a Keyence machine under high magnification after every 24 milled surfaces to make sure there was no visible tool wear.

The milling acoustic signals are collected in real time by the microphone placed inside the milling chamber. Each milled surface consists of multiple tool paths, which include air cutting and changes in cutting direction. The acoustic sensor starts recording before the milling process is initiated. Therefore, acoustic data outside the stable zone was also recorded. However, this data might create an issue for ML training as air cutting and changes in cutting direction

are not related to the machined surface in the stable milling zone. For precise ML training data, acoustic signals of air cutting were excluded, as shown in Fig. 6. Acoustic signals corresponding to the changes in cutting direction were also identified and excluded in preparing the dataset for ML training and testing. The acoustic signals consist of the signal from each cutting path inside the stable zone (P1 and P2 in Fig. 6a) and the signal from the step-over segment when the cutting direction changes.

2.3 Surface data measurement and acoustic data post-processing

Surface topographies of the milled surfaces inside the stable milling zone were measured using the Keyence VR-3000 system. Figure 7 shows two examples of the measured surface topography under different milling conditions. The surface topographies were converted and output as height matrices, where the column and row of such matrices represent the pixel position. The value inside each matrix cell represents the measured height of the surface at that pixel location. Then, the surface data are saved in CSV format for subsequent use in training ML models. Also, the surface roughness parameters can be calculated from this surface data, such as 3D surface roughness S_a , maximum height S_z , root mean square S_q , skewness S_{sk} , and kurtosis S_{ku} . The detailed description of surface roughness parameters is described in Sect. 4.1.

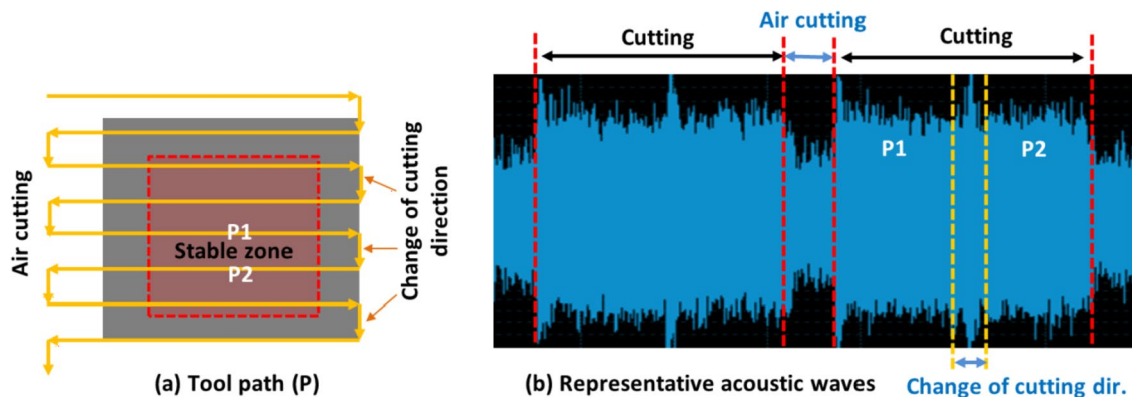
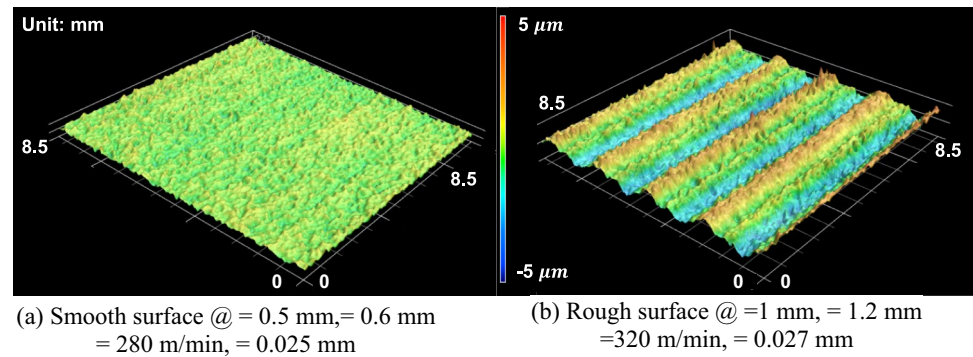
**Fig. 6** Acoustic data collection

Fig. 7 Representative surface topography

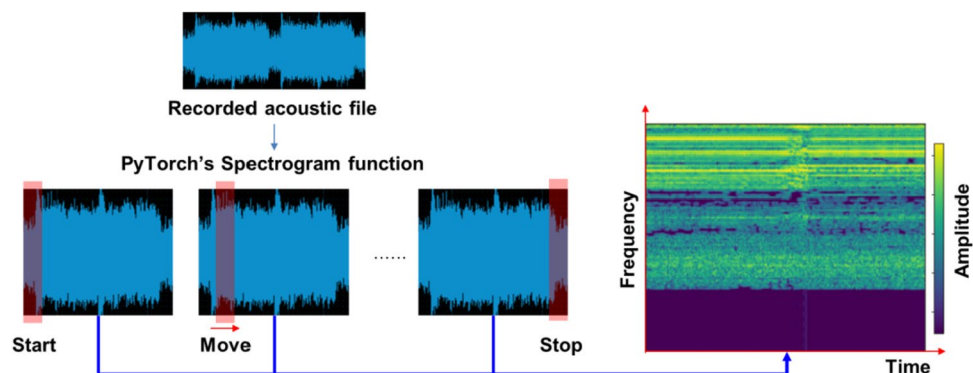
Acoustic signals are commonly used to identify milling stabilities [49] and chatter [50]. In contrast to using direct learning on the raw waveforms in the previous ML work for audio [51, 52], the audio signals are used to create a time–frequency map from the power magnitude obtained through Short-time Fourier Transform (STFT), which captures the changes over time of each frequency component. Figure 8 shows the process of the spectrogram conversion. It converts the time-series data into a spectrogram, revealing the frequency and amplitude change with time. First, a small-time step was used to slice the time-series sound signal (red box region) and use this slice to perform Fourier Transform (FT) to get frequency-amplitude information. Then, the FT process repeats at the next time step. After all time steps were processed, all frequency-amplitude data was stitched with the time stamp to get the whole 3-dimensional spectrum representing the acoustic data. The X- and Y-axis are the time stamp and frequency, respectively, and the color legend represents the amplitude. Note the shape of the spectrums is different as the time of each milling process is different.

3 Multimodal data-based ML models

The proposed ML model aims to use acoustic signals and milling parameters as the model inputs to predict surface topography. Generative Adversarial Network (GAN) is one of the popular deep learning architectures used for image generation. Such architecture trains two neural networks (i.e., Generator and Discriminator) in a zero-sum game until the Discriminator cannot tell the true images. The layer structure of a GAN model [53] for efficient ML training and inference has been adopted in this study. However, a key change was introduced that the traditional loss function, e.g., cycle consistency loss, is bypassed since this study focuses on mapping milling parameters and acoustic signals to surface topography. This alteration leverages the selected GAN model to align more closely with the specific requirements of this study. The ML model in this work is built using PyTorch, one of the most widely used open-source deep learning libraries.

3.1 ML problem formulation

In the presented study, the prediction of milled surface topography is formulated as a supervised ML problem. In general, the milling parameters and milling acoustic signals are treated as the input information for the ML model,

Fig. 8 Convert acoustic data into spectrum

and the surface topography is the ground truth for the predictions.

Let x_p denotes the input milling parameters, x_s denotes the milling acoustic signals, $\tilde{y}$ denotes the predicted surface topography, and y denotes the measured surface map (surface position X, Y, and height Z). A CNN called a generator, denoted as $G(\cdot)$ can be used to map the input information to the corresponding surface topography, where $\mathcal{L}$ denotes a loss function, and $\|\cdot\|_F$ represents the Frobenius norm. Hence, the supervised ML problem can be formulated as follows:

$$\min_G \mathcal{L}(\tilde{y}, y) = \min_G \|\tilde{y} - y\|_F^2 = \min_G \|G(x_p, x_s) - y\|_F^2 \quad (1)$$

Solving Eq. (1) has several challenges, as discussed in the following sub-sections.

3.2 Fusing inputs from different modalities

Challenge 1: Input signals x_p and x_s have different modalities. The milling parameters x_p are represented as 4-dimensional vectors, while the milling audio data x_s are represented as acoustic signal waves. Hence, these two input signals have different representations and statistical distributions. As a result, directly using these signals as inputs for the generator $G(\cdot)$ can lead to poor results because of the generator's inability to fuse the two inputs.

Approach 1: To address Challenge 1, we propose to use a trainable Encoder $E(\cdot)$ that includes several transformations on both x_p and x_s , and a trainable linear layer to process these two inputs that come from different modalities. The Encoder

$E(\cdot)$ fuses the two separated input signals x_p and x_s , into an input tensor x that encodes both the milling parameters and milling acoustic signals: $E(x_p, x_s)$. Hence, the input to $G(\cdot)$ is no longer x_p and x_s , but the fused input x .

Figure 9 shows the encoder process in detail. To ensure size consistency, both the model input tensor and output topography are set to 256×256 , which matches the dimension of the measured surface topographies. Importantly, input tensors accommodate multiple channels to encode additional data. The tensor is transformed from the 4×1 milling parameters tensor into a 256×1 tensor using linear layers and then augmented for dimension matching. Linear transform is a common machine learning technique involving mapping low-dimensional data into higher dimensions [36]. Given the milling parameters $x_p = [a_e \ a_p \ f_z \ v]^T$ and the linear layer noted as $y = Wx + B$, the transform of the 4×1 tensor into a 256×1 tensor can be written as:

$$x_p^{expand} = W \cdot x_p + B. \quad (2)$$

In which $W = [w_{ij}]$, where $i = 1, 2, 3 \dots 256$ and $j = 1, 2, 3, 4$; $B = [b_1, b_2 \dots, b_{256}]^T$. W and B represent the weights and bias of the linear transform and can be updated by the training loss. Then, the 256×1 tensor duplicates itself to match the 256×256 dimension.

In order to fuse the acoustic signals with the milling parameters, the acoustic signals are first transformed into a spectrogram, which is then resized into the same shape as the milling parameters tensor using PyTorch's resize function (bilinear interpolation). Finally, the milling parameters and spectrograms are concatenated into a 2-channel tensor

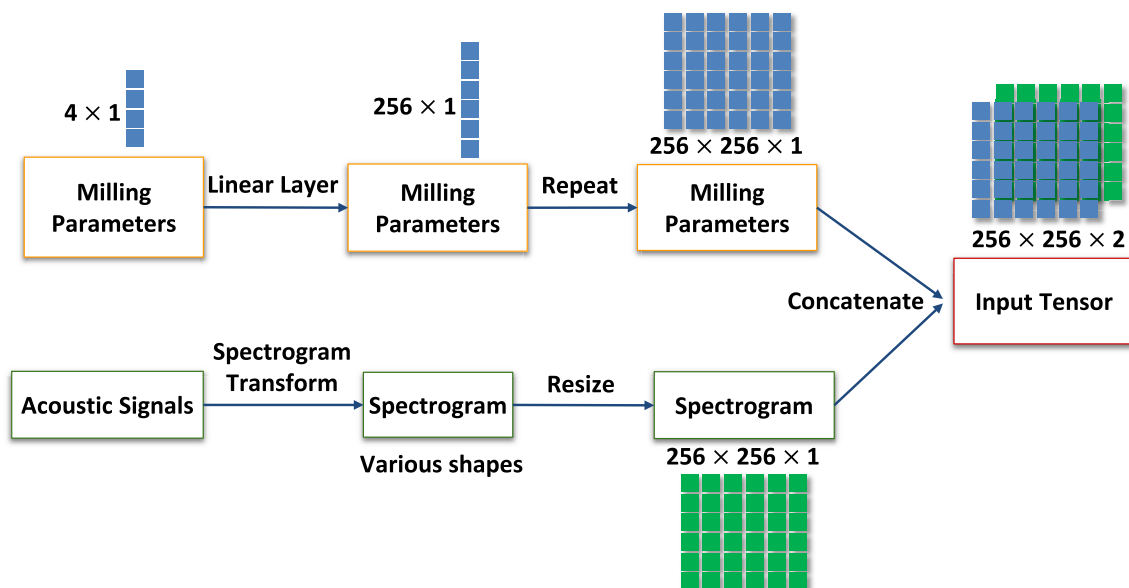


Fig. 9 Fusing the two inputs from different modalities using the encoder

with dimensions $256 \times 256 \times 2$. Concatenation is often used to encode additional data and can be applied to different types of tensors, including videos with audio [54] and numbers with images [55]. Several other encoders are also evaluated, and their performance is reviewed in Sect. 4.2. It should be pointed out that there are many alternative approaches to encoding data. It could be further investigated in another study.

3.3 Learning the variations in the surface topography using GANs

Challenge 2: While loss function $\mathcal{L}$ in Eq. (1) measures pixel difference, it might not be sensitive enough to capture all the intricate details, patterns, and imperfections across the surface topographies. As y is not the unique mapping of x_s and x_p , other variations of y can also correspond to the same surface topography. Hence, the model cannot learn the variations in the surface topography only by minimizing the norm difference between the prediction $\tilde{y}$ and the ground truth y . Furthermore, since the training dataset has a limited number of data points (less than 1000), the overparameterized $G(\cdot)$ will overfit the training data and thus will not generalize to new unseen data.

Approach 2: To address Challenge 2, the GAN framework [47, 53] is proposed to effectively train the generator $G(\cdot)$ to predict surface topography. The GAN learning framework consists of two separate DNNs: the generator $G(\cdot)$ and the discriminator $D(\cdot)$. In this problem, the goal of the generator is to generate realistic and correct surface topography $\tilde{y}$ from the input milling information x , while the discriminator's goal is to classify a given surface topography as generate (synthesized by the G network), or real (actual measurement of the surface). The

two networks have opposing optimization goals. Hence G and D play the following minimax game:

$$\min_G \max_D \log D(y) + \log(1 - D(G(E(x_p, x_s))))). \quad (3)$$

Optimizing the discriminator:

The discriminator can be trained using the following optimization function:

$$\min_D \mathcal{L}_D = -[\log D(y) + \log(1 - D(G(E(x_p, x_s))))]. \quad (4)$$

Several augmentation strategies have been investigated, including cropping, flipping, scaling, and blurring the input images to create a more diverse training set. As a result, the input images have a wide range of variations in the surface topography. Hence, the discriminator must learn the actual characteristics of the surface rather than memorize the measured images in the dataset.

Optimizing the encoder and generator:

$$\min_{E,D} \mathcal{L}_G = \underbrace{-\log D(G(E(x_p, x_s)))}_{\text{adversarial loss}} + \underbrace{\|\beta \cdot G(E(x_p, x_s)) - y\|_F^2}_{\text{correctness loss}}. \quad (5)$$

In Eq. (5), β is the balancing hyper-parameters between adversarial loss and correctness loss. The Generator network has two goals: (1) synthesize realistic-looking surface images that fool the discriminator into thinking those images are actual measurements. (2) Synthesize images that have correct surface topography. The two goals are reflected by the adversarial loss term and correctness loss term in (5), respectively. The details of the overall training procedure are summarized in Fig. 10 and Algorithm 1 (Pseudo Code). The algorithm delineates the model's training process in structured steps: Inputs and Outputs

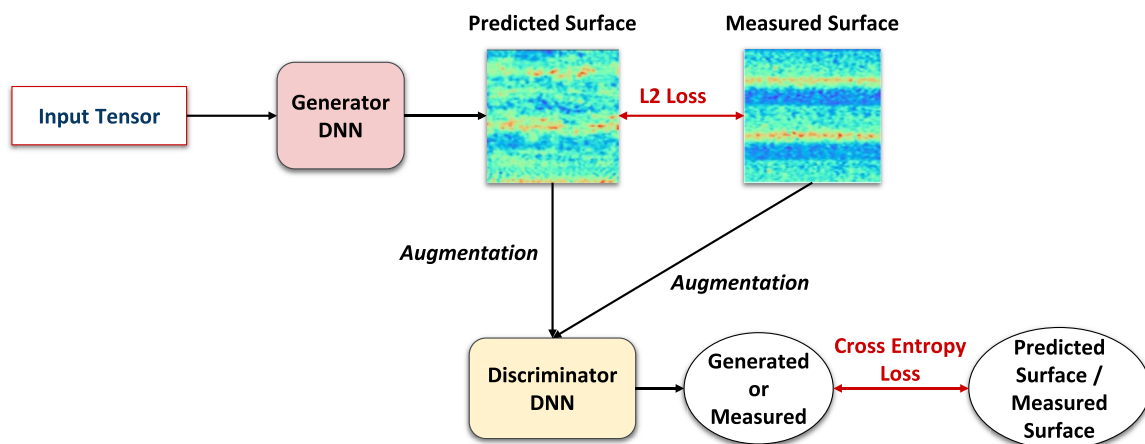


Fig. 10 The overall training procedure

(line 1–2): the algorithm starts by defining inputs: dataset D , which contains milling parameters x_p , milling acoustic signals x_s , measured surface topography y . Additionally, it also requires a learning rate α to guide the optimization and a balancing factor β to weigh the different components of the loss function. Line 2 defines the final outputs after training as the trained encoder (E) and trained generator (G). Initialization (line 3): The weights of the E , discriminator (D), and G are initialized with random values. This random initialization serves as the starting point for the optimization process. Training Loop (lines 4–15): line 4 defines a loop that iterates over each data tuple from D , which encompasses the x_p , x_s and corresponding measured surface topography (y). Discriminator Training (lines 5–10): line 5 is a comment indicating the following lines are for discriminator training. In line 6, the surface topography ($y_{generate}$) is generated using the E and D given the input x_p and x_s . Both the generated ($y_{generate}$) and the measured topography (y) undergo data augmentation in line 7. Data augmentation is a technique to artificially expand the dataset by creating slightly modified copies of existing data. It enhances the robustness and generalization capability of the model. In line 8, the D assesses the

topographies, ideally scoring real topographies near one and generated ones near 0. A loss for D is calculated using the scores (line 9) and is utilized to update the weights of this loss and the learning rate (α) at line 10. Encoder and Generator Training (lines 11–15): line 11 is a comment indicating the following lines are for training E and G . The surface topography ($y_{generate}$) is generated the same as line 6 (line 12). This generated topography is evaluated by the D to give a prediction score at line 13. In line 14, the loss for both E and G is computed in two parts: 1. Cross-entropy between discriminator's (D) prediction for the generated topography and an ideal score of 1. 2. Mean Squared Error (MSE) between the generated topography and the measured one. The balancing factor (β) weighs the importance of this accuracy term in the loss function. Finally, in line 15, the combined loss updates weights for both E and G with the given learning rate (α). Figure 10 shows the overall training procedure: the generator predicts the surface to compare with the measured surface and get L_2 loss. The predicted and measured surfaces will be fed to the Discriminator to be classified as generated or measured and get CSE (cross-entropy loss). Both losses are used to update the model weights.

Algorithm 1: Multimodal Data-Based ML Algorithm for Prognosis of Surface Topography

```

1. Input: Dataset  $D$  consists of milling parameters  $x_p$ , milling acoustic signals  $x_s$ , measured
   surface topography  $y$ . Learning rate  $\alpha$  and balancing factor  $\beta$ .
2. Output: Trained Encoder  $E$ , trained generator  $G$ .
3.  $E, D, G \leftarrow \text{random.init}(E, D, G)$  #random initialize weights
4. for  $(x_p, x_s, y)$  in  $d$  do:
5.     #Train Discriminator ( $D$ )
6.      $y_{generate} \leftarrow G(E(x_p, x_s))$ 
7.      $y_{generate}, y \leftarrow \text{augment}(y_{generate}), \text{augment}(y)$ 
8.      $\text{pred}_{generate}, \text{pred}_{real} \leftarrow D(\text{generate}), D(y)$ 
9.      $\text{loss} \leftarrow \text{cross.entropy}(\text{pred}_{generate}, 0) + \text{cross.entropy}(\text{pred}_{real}, 1)$ 
10.     $\text{update}(Dis, \text{loss}, \alpha)$ 
11.    #Train Encoder  $E$  and Generator  $G$ 
12.     $y_{generate} \leftarrow G(E(x_p, x_s))$ 
13.     $\text{pred}_{generate} \leftarrow D(y_{generate})$ 
14.     $\text{loss} \leftarrow \text{cross.entropy}(\text{pred}_{generate}, 1) + \beta \cdot \text{MSE}(y_{generate}, y)$ 
15.     $\text{update}(E, \text{loss}, \alpha), \text{update}(G, \text{loss}, \alpha)$ 

```

3.4 ML experiment setup

The dataset has a total of 1296 data points. 80% of the data were randomly selected for the training set and the rest for the validation set. To enlarge the training set, one or more augmentation techniques are randomly applied: random cropping, horizontal flipping, scaling, and Gaussian blurring. During the validation phase, no augmentation technique is used, and all validation images are resized to 256×256 .

The architectures proposed in [48] were used for the discriminator and generator model architectures. Specifically, the *ResNetGenerator* with a depth of 5 blocks was used for the Generator, and the *NLayer Discriminator* with a depth of 2 layers was used for the Discriminator. A linear layer of 4 input neurons and 256 output neurons was used for the encoder. Adam optimizer with a learning rate of 0.0003 and weight decay of 0.0001 was used to train the three models for a total of 1000 epochs. Here, “epochs” refers to complete cycles through the entire training dataset. One epoch consists of multiple training steps to cover every sample once in the dataset. Each training step is an individual iteration where the model updates its weights based on a subset (batch) of the training dataset. The balancing factor was set to $\beta = 1$.

4 Results and discussions

4.1 ML model predictions

The ML models were trained in two ways: (1) with only four milling parameters as the input to generate surface topography and (2) integrating four milling parameters and acoustic signal spectrum as input to generate surface topography. Both types of training were conducted using the same training method described in Sect. 3. In this section, we plot the training results as color maps to intuitively show the difference between the measurements and the training results.

Then, the predicted surfaces are analyzed quantitatively using surface roughness metrics.

Figure 11 shows the top view of a few examples of the ML model’s training results at the end of the training process. Each subfigure contains 28 surfaces ($7\text{-row} \times 4\text{-column}$ surface matrix), and the color in each map indicates the height of the surface. Figure 11a shows the measured surface topographies as the ground truth. Figure 11b shows the outputs from the ML model trained with the multimodal data (i.e., integrated milling parameters and acoustic signals) as the input. Figure 11c is the output from benchmark training, where only milling parameters were used as the input. The images show that the ML model can predict similar surface data distributions to those of the measured surfaces. At the same time, the benchmark training results cannot learn such distributions, and the outputs are not similar to the measurement.

Figure 12 shows the measured surface compared with the model output in a three-dimensional view. The proposed method generates the multimodal prediction using the corresponding milling parameters and acoustic signals from the validation dataset. The benchmark prediction is generated using only corresponding milling parameters. In Fig. 12, the top left measured surface has three noticeable feed marks. This feature is also shown on the multimodal predicted surface topography but is not on the benchmark training. The top right image of Fig. 12 shows a relatively smoother surface. The multimodal prediction also has a similarly smooth surface, and the benchmark prediction is almost the same as the left side. Thus, the three-dimensional topography surfaces intuitively suggest the proposed method can learn the hidden relation between the actual surfaces and the multimodal input data (e.g., milling parameters and acoustic signals) but not the benchmark data. Since the benchmark model only takes milling parameters as input, it lacks sufficient information to learn the hidden relationship between inputs and surface topography. Consequently, the benchmark predictions are trivial and fail to match the measurements. In fact, the

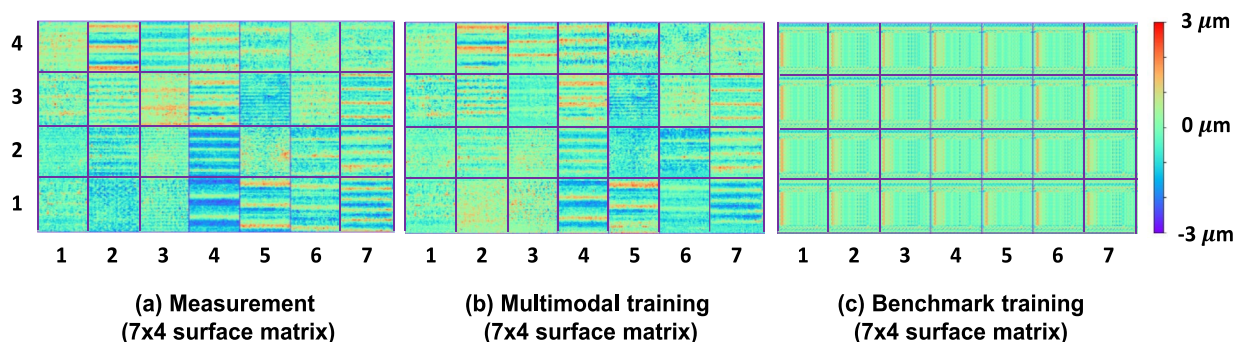


Fig. 11 Surface topographies by measurement, ML model, and benchmark

Fig. 12 Examples of ML predicted surface topography compared with measured ones

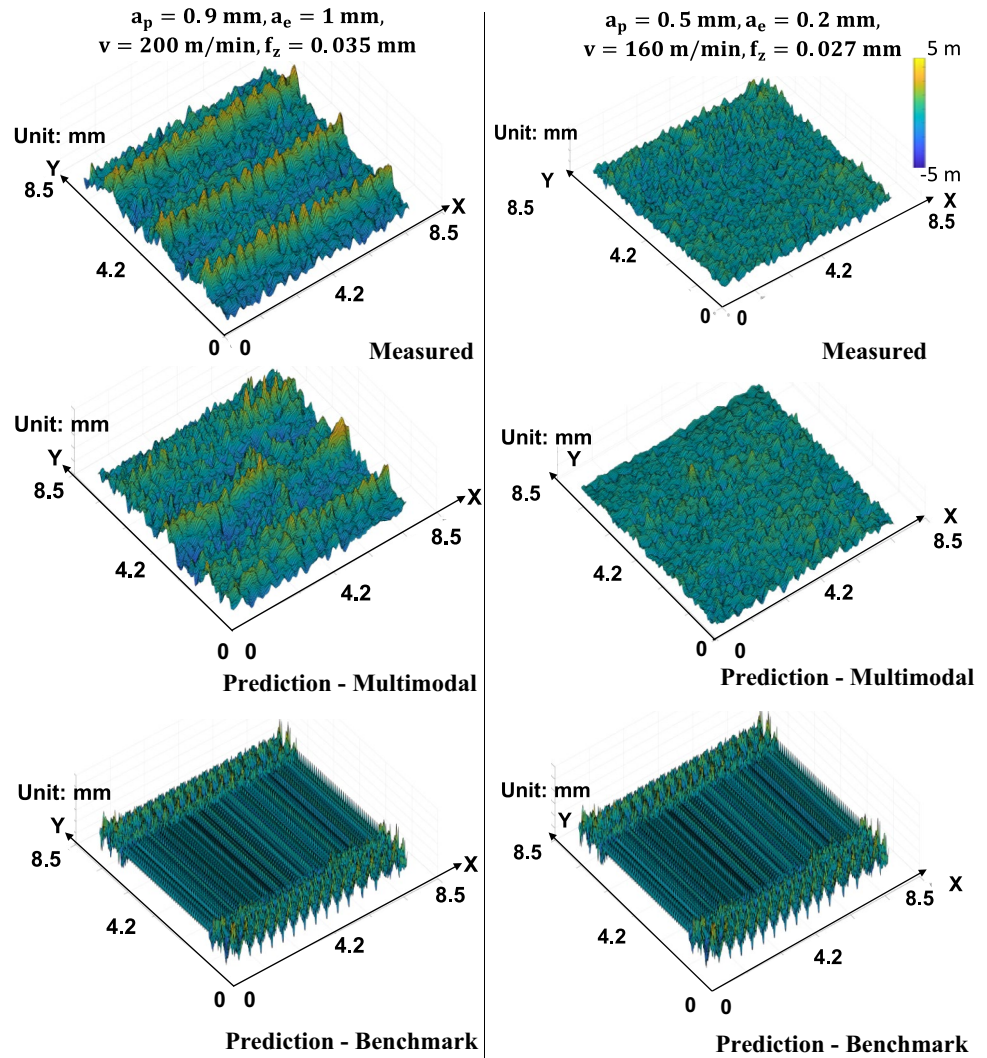


Table 4 Surface indicators of validation dataset (260 surfaces)

Methods		Surface indicators (averaged)				
		S_a (μ m) Arithmetical mean height	S_z (μ m) Maximum height	S_q (μ m) Root mean square height	S_{sk} Skewness	S_{ku} Kurtosis
		$\frac{1}{A} \iint Z dx dy$	$ Max Peak + Min Valley $	$\sqrt{\frac{1}{A} \iint Z^2 dx dy}$	$\frac{1}{S_q^3} [\frac{1}{A} \iint Z^3 dx dy]$	$\frac{1}{S_q^4} [\frac{1}{A} \iint Z^4 dx dy]$
1	Measured	0.649	6.897	0.816	0.198	3.290
2	Multimodal	0.504	6.538	0.654	0.278	3.801
3	Δ_1	0.145	0.359	0.162	-0.080	-0.511
4	Benchmark	0.751	14.099	1.072	0.488	7.262
5	Δ_2	-0.102	-7.202	-0.256	-0.290	-3.972

benchmark predictions are almost identical for every input and consistently show rough regions at the front and back of the surface.

The outputs of the proposed model are surface height maps (positions X, Y, and height Z) of the milled surfaces instead of images. The general quantitative measures of image correlation, such as the sum of squared differences

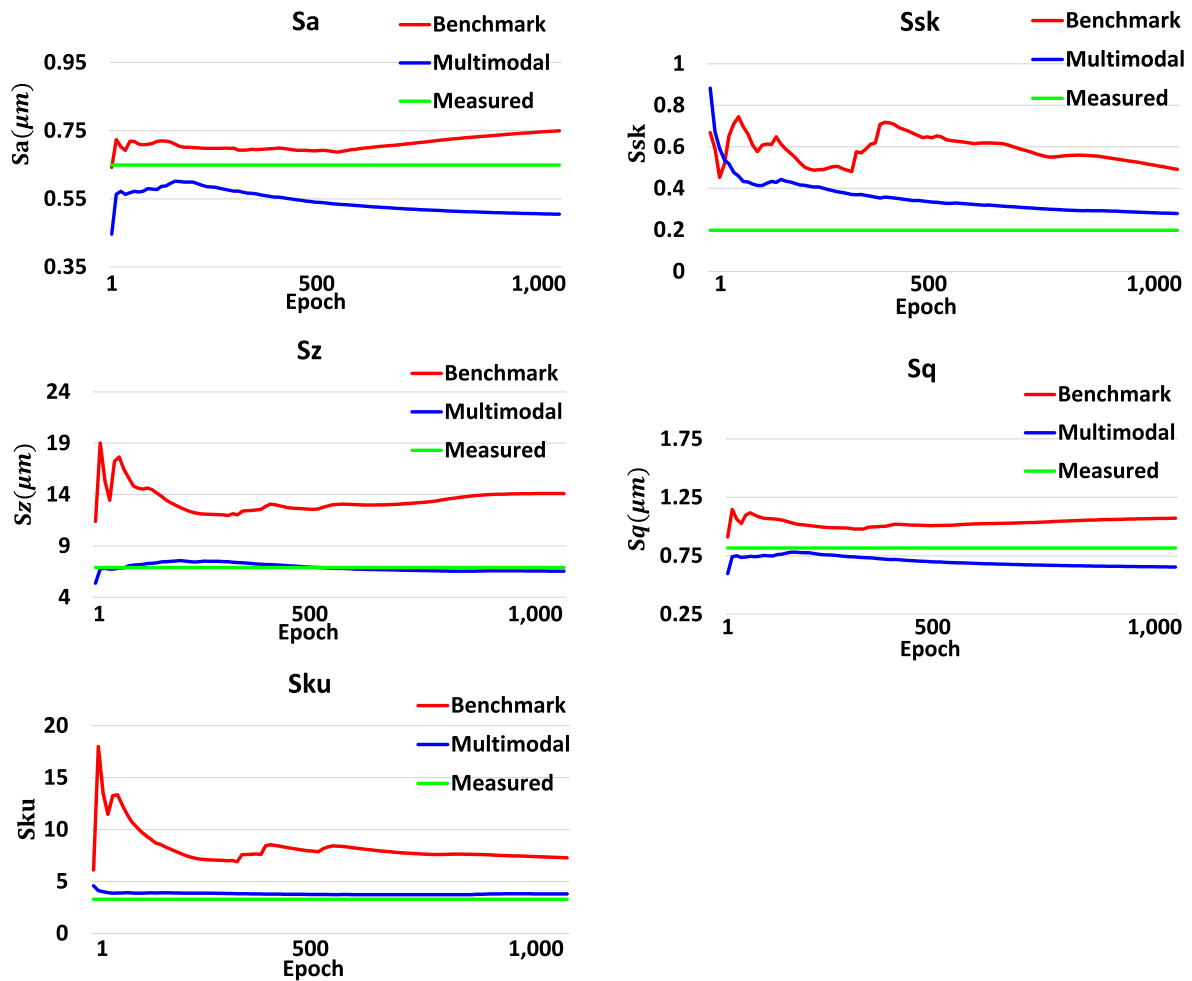


Fig. 13 Surface topography parameters testing errors with epochs

(SSD) and cross-correlation (CC), are widely used for image analysis but are not suitable for analyzing surface topography. The proposed method uses the ISO surface parameters (Table 4) to analyze the predicted surfaces' deviation quantitatively. Note the surface indicators are calculated as the average of the 260 surfaces in the validation dataset. The 3D surface roughness S_a expresses the absolute value of the difference between the height at each surface point and the theoretical perfect flat surface (arithmetical mean of the surface). S_z defines the sum of the highest peak value and lowest valley value. S_q (root mean square or RMS) is the standard deviation of the heights in the area. S_{sk} (skewness) values the degree of bias of the roughness shape, and S_{ku} (kurtosis) represents the sharpness of the roughness profile. The variables x, y stand for the pixel position on the surface, while Z represents the relative difference measured from the surface mean plane. Thus, each measured surface topography has a set of five surface parameters to define the surface geometrical properties. The model's performance was evaluated on the validation dataset after the training

phase, which consisted of previously unseen data. As delineated in Table 4, the proposed multimodal approach generally outperforms the benchmark method except for the S_a value. This suggests that while the benchmark method is more adept at capturing the pattern of surface mean height, the multimodal method excels in other aspects. Δ_1 stands for the difference between the measured value and multimodal prediction, while Δ_2 denotes the difference between the measured value and benchmark prediction.

Figure 13 presents the surface indicators of both measured and predicted values throughout the training epochs. The S_z , S_q , and S_{ku} subfigures suggest that the multimodal method offers superior surface sharpness and deviation prediction than the benchmark as the multimodal prediction aligns closely with the measured values early in training, maintaining minor discrepancies throughout. In the S_{sk} subfigure, the multimodal prediction converges to a value closer to the measurements, whereas the benchmark exhibits consistent fluctuations throughout the training. This indicates the multimodal's enhanced capability in predicting the

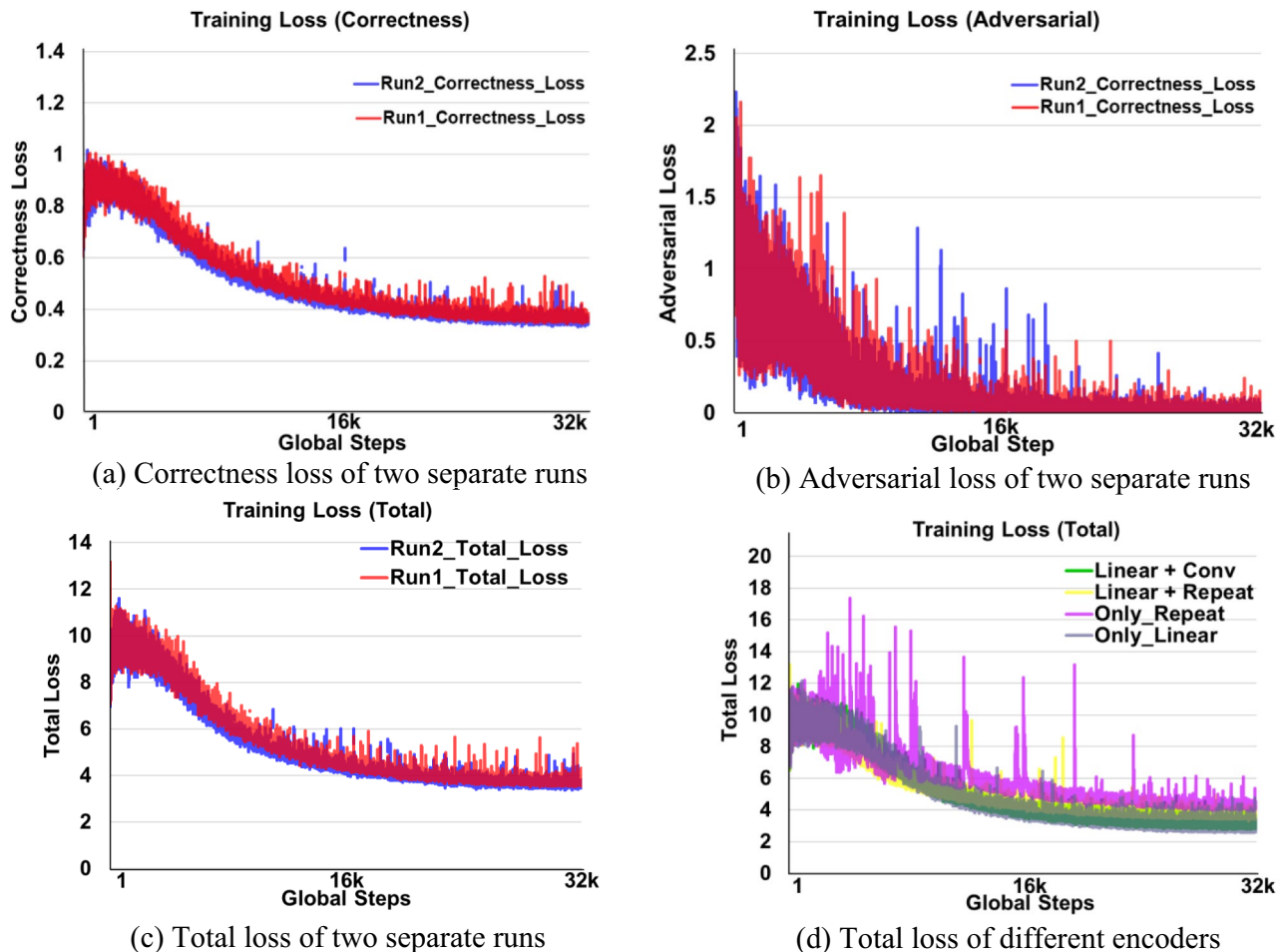


Fig. 14 Training loss of separate runs against global steps

roughness shape bias over the benchmark. The Sa subfigure suggests that the multimodal predicts a relatively smoother surface compared to both the benchmark and the measured surface.

4.2 ML results analysis

From the ML perspective, the model successfully captured the relationship between the multimodal input data and the corresponding surface topography during the training phase. The predicted surface topography is almost identical to the actual measurement by the end of the training loop. Figure 14 (a) correctness loss, (b) adversarial loss, and (c) total loss shows the two separate runs of the training loss (referring to Eq. 5) of the multimodal training process to verify the proposed model's stability. Run1 and Run2 have randomly and independently selected the training dataset. The X-axis, global steps, stands for the cumulative count of the individual training step iteration. As the figures suggest, all three losses steadily decrease along the global steps and converge

around 25 k steps. Additionally, with different training datasets, Run1 and Run2 show almost the same trend on the loss trend, which indicates the model is stable against different data. Figure 14d illustrates the training loss of different encoders. “Linear + Repeat” represents the training loss for the encoder used in this study. The “Linear + Conv” is similar to that of the proposed encoder, which uses a linear layer to convert the 4×1 milling parameters tensor into a 256×1 tensor, followed by a convolution layer to expand it to a 256×256 tensor. Two other encoding approaches are also compared: “Only_Repeat” and “Only_Linear.” It's worth noting that the “Only_Linear” encoder has also been utilized in literature [36]. The former expands the 4×1 tensor to a 256×256 tensor solely via repetition, while the latter achieves this expansion using only a linear layer.

From the results in Fig. 14d, it's clear that the “Only_Repeat” encoder records the highest loss (4.274) at the end of the training cycle. This is approximately 1.33, 1.39, and 1.16 times higher than the losses of the “Only_Linear” encoder (3.030), the “Linear + Conv” encoder (3.077), and

Fig. 15 Training loss of different depth

the “Linear + Repeat” encoder (3.689), respectively. It is noted that the computational complexity, a measure of the computing resources (time and memory) required by an algorithm during its execution, is also a crucial factor in evaluating these encoders. The “Linear + Conv” and “Only_Linear” encoders have the highest computational complexity among the four models. This is because both the linear transformation and convolution operations involve 256×256 tensor multiplication, which is computationally intensive. The “Linear + Repeat” encoder has slightly lower computational complexity, given that duplication is less complex than either linear transformation or convolution. The “Only_Repeat” encoder has the least computational complexity due to its straightforward repeating operation. Based on the tradeoff between training loss and computational complexity, the “Linear + Repeat” encoder was used in this study.

Multiple Generator and Discriminator architectures were tested with varying depths, and it is important to note that the deeper architectures did not perform the task better, while shallower architectures produced low-quality results. Figure 15 displays the training loss of the Generators and Discriminators of different depths throughout the training process. The notation used indicates the depth of the architecture. For example, “G4D2” represents a depth of 4 blocks of *ResNetGenerator* for the Generator and a depth of 2 layers of *NLayer Discriminator* for the Discriminator. Notably, G5D2 exhibited the best performance, with the training loss converging and stabilizing around 10 k global steps.

5 Conclusions

This paper proposes a multimodal data-driven ML model integrating different milling process information to predict machined surface topography. The full-factorial design of experiments generates a dataset to cover various milling conditions. The model’s training was initially confined to milling parameters alone to set a benchmark. In subsequent training phases, a fusion of both milling parameters and real-time processed acoustic data gathered during milling operations was employed. Integrating real-time acoustic data into the model showcased a discernible enhancement in the prediction accuracy. The comprehensive comparisons between predicted and measured surface topographies solidified this observation. This advancement underscores the pivotal role of assimilating process physics data into ML training, a strategy that brings efficiency and heightened accuracy to the predictive modeling of milling processes.

The choice of employing the Generative Adversarial Network (GAN) model was influenced by its distinctive attributes. Foremost among these is GAN’s proficiency in managing high-dimensional datasets, outpacing traditional algorithms like Random Forests or Support Vector Machines. Additionally, GANs have shown a remarkable ability to decode the intricate relationships between various milling parameters and the resultant acoustic emissions, an accomplishment achieved due to their natural inclination towards recognizing non-linear associations. This quality is paramount in this study, which deals with the complex and nuanced prediction of milled surface topographies. Notably, the GAN’s unsupervised learning nature makes it uniquely positioned to handle the multifaceted relationships between milling parameters and acoustic

data, sidestepping the need for intricate feature engineering required by other Neural Networks.

In summary, as a complement to traditional measurement techniques, the proposed approach promises enhanced precision and efficiency, significantly reducing associated measurement time and costs. While our model excels within the confines of our dataset, we recognize the inherent challenges of extrapolation and believe it presents an intriguing avenue for future research.

Acknowledgements The authors would like to thank the National Science Foundation for the financial support of this study under Grant CMMI-2040358/2040288.

Author contributions LH: data curation (lead), formal analysis (lead), investigation (lead), methodology (lead), software (equal), validation (lead), visualization (lead), writing—original draft (lead). HP: data curation (support), formal analysis (equal), investigation (equal), methodology (support), software (lead), validation (support), visualization (equal), writing—original draft (equal). SS: data curation (support), methodology (support). CC: investigation (support), methodology (support), validation (support), writing—original draft (support). JZ: formal analysis (support), investigation (support), visualization (support), writing—original draft (support). BY: conceptualization (support), methodology (support), project administration (support), software (support), supervision (support), writing—review and editing (equal). RG: conceptualization (support), funding acquisition (equal), project administration (support), supervision (support), writing—review and editing (equal). YBG: conceptualization (lead), funding acquisition (lead), project administration (lead), resources (lead), supervision (lead), writing—review and editing (lead).

Funding This work has been supported by the National Science Foundation under Grant CMMI-2040358/2040288.

Data availability statement The data that support the findings of this study are available from the corresponding author upon request.

Declarations

Conflict of interest The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

- Maleki E, Bagherifard S, Bandini M, Guagliano M (2021) Surface post-treatments for metal additive manufacturing: progress, challenges, and opportunities. *Addit Manuf* 37:101619
- Koshy P, Dewes R, Aspinwall D (2002) High speed end milling of hardened AISI D2 tool steel (~ 58 HRC). *J Mater Process Technol* 127(2):266–273
- Sun J, Guo Y (2009) A comprehensive experimental study on surface integrity by end milling Ti–6Al–4V. *J Mater Process Technol* 209(8):4036–4042
- Shokrani A, Dhokia V, Newman ST (2016) Investigation of the effects of cryogenic machining on surface integrity in CNC end milling of Ti–6Al–4V titanium alloy. *J Manuf Process* 21:172–179
- Jebaraj M, Kumar P, Anburaj R (2020) Effect of LN₂ and CO₂ coolants in milling of 55NiCrMoV7 steel. *J Manuf Process* 53:318–327
- Nouri M, Fussell BK, Ziniti BL, Linder E (2015) Real-time tool wear monitoring in milling using a cutting condition independent method. *Int J Mach Tools Manuf* 89:1–13
- Altintas Y, Stépán G, Merdol D, Dombóvári Z (2008) Chatter stability of milling in frequency and discrete time domain. *CIRP J Manuf Sci Technol* 1(1):35–44
- 25178-2:2021 I, Geometrical product specifications (GPS)—surface texture: areal—part 2: terms, definitions and surface texture parameters
- Novovic D, Dewes R, Aspinwall D, Voice W, Bowen P (2004) The effect of machined topography and integrity on fatigue life. *Int J Mach Tools Manuf* 44(2–3):125–134
- Salahshoor M, Li C, Liu Z, Fang X, Guo Y (2018) Surface integrity and corrosion performance of biomedical magnesium-calcium alloy processed by hybrid dry cutting-finish burnishing. *J Mech Behav Biomed Mater* 78:246–253
- Hashimoto F, Guo Y, Warren A (2006) Surface integrity difference between hard turned and ground surfaces and its impact on fatigue life. *CIRP Ann* 55(1):81–84
- Cooper LF (2000) A role for surface topography in creating and maintaining bone at titanium endosseous implants. *J Prosthet Dent* 84(5):522–534
- Anonymous (2022) Toyota's Nov global vehicle production rises 1.5% to record 833,104. *Reuters News*
- Montgomery D, Altintas Y (1991) Mechanism of cutting force and surface generation in dynamic milling. *J Eng Ind* 113(2):160–168
- Amin AN, Ismail AF, Khairushima MN (2007) Effectiveness of uncoated WC–Co and PCD inserts in end milling of titanium alloy—Ti–6Al–4V. *J Mater Process Technol* 192:147–158
- Altintas Y, Budak E (1995) Analytical prediction of stability lobes in milling. *CIRP Ann* 44(1):357–362
- Schmitz TL, Couey J, Marsh E, Mauntler N, Hughes D (2007) Runout effects in milling: surface finish, surface location error, and stability. *Int J Mach Tools Manuf* 47(5):841–851
- Arizmendi M, Campa F, Fernández J, De Lacalle LL, Gil A, Bilbao E, Veiga F, Lamikiz A (2009) Model for surface topography prediction in peripheral milling considering tool vibration. *CIRP Ann* 58(1):93–96
- Ding Y, Zhu L, Zhang X, Ding H (2010) A full-discretization method for prediction of milling stability. *Int J Mach Tools Manuf* 50(5):502–509
- Altintas Y, Jin X (2011) Mechanics of micro-milling with round edge tools. *CIRP Ann* 60(1):77–80
- Costes JP, Moreau V (2011) Surface roughness prediction in milling based on tool displacements. *J Manuf Process* 13(2):133–140
- Kaymakci M, Kilic Z, Altintas Y (2012) Unified cutting force model for turning, boring, drilling and milling operations. *Int J Mach Tools Manuf* 54:34–45
- Rao B, Dandekar CR, Shin YC (2011) An experimental and numerical study on the face milling of Ti–6Al–4V alloy: tool performance and surface integrity. *J Mater Process Technol* 211(2):294–304
- Jin X, Altintas Y (2012) Prediction of micro-milling forces with finite element method. *J Mater Process Technol* 212(3):542–552
- Wu H, Zhang S (2014) 3D FEM simulation of milling process for titanium alloy Ti6Al4V. *Int J Adv Manuf Technol* 71(5):1319–1326
- Felhoc C, Karpuschewski B, Kundrák J (2015) Surface roughness modelling in face milling. *Procedia CIRP* 31:136–141
- Suresh Kumar Reddy N, Venkateswara Rao P (2005) Selection of optimum tool geometry and cutting conditions using a surface roughness prediction model for end milling. *Int J Adv Manuf Technol* 26(11):1202–1210
- Çolak O, Kurbanoğlu C, Kayacan MC (2007) Milling surface roughness prediction using evolutionary programming methods. *Mater Des* 28(2):657–666

29. Yang D, Liu Z (2015) Surface topography analysis and cutting parameters optimization for peripheral milling titanium alloy Ti–6Al–4V. *Int J Refract Metal Hard Mater* 51:192–200
30. Karandikar J, Honeycutt A, Schmitz T, Smith S (2020) Stability boundary and optimal operating parameter identification in milling using Bayesian learning. *J Manuf Process* 56:1252–1262
31. Karandikar J, Honeycutt A, Smith S, Schmitz T (2020) Milling stability identification using Bayesian machine learning. *Procedia CIRP* 93:1423–1428
32. Zain AM, Haron H, Sharif S (2010) Prediction of surface roughness in the end milling machining using Artificial Neural Network. *Expert Syst Appl* 37(2):1755–1768
33. Peng C, Wang L, Liao TW (2015) A new method for the prediction of chatter stability lobes based on dynamic cutting force simulation model and support vector machine. *J Sound Vib* 354:118–131
34. Paturi UMR, Cheruku S, Pasunuri VPK, Salike S, Reddy N, Cheruku S (2021) Machine learning and statistical approach in modeling and optimization of surface roughness in wire electrical discharge machining. *Mach Learn Appl* 6:100099
35. Liu M, Cheung CF, Senin N, Wang S, Su R, Leach R (2020) On-machine surface defect detection using light scattering and deep learning. *JOSA A* 37(9):B53–B59
36. Cao L, Huang T, Zhang X-M, Ding H (2020) Generative adversarial network for prediction of workpiece surface topography in machining stage. *IEEE/ASME Trans Mechatron* 26(1):480–490
37. Rahimi MH, Huynh HN, Altintas Y (2021) On-line chatter detection in milling with hybrid machine learning and physics-based model. *CIRP J Manuf Sci Technol* 35:25–40
38. Griffin JM, Diaz F, Geerling E, Clasing M, Ponce V, Taylor C, Turner S, Michael EA, Mena FP, Bronfman L (2017) Control of deviations and prediction of surface roughness from micro machining of THz waveguides using acoustic emission signals. *Mech Syst Signal Process* 85:1020–1034
39. Gauder D, Biehler M, Gözl J, Schulze V, Lanza G (2022) In-process acoustic pore detection in milling using deep learning. *CIRP J Manuf Sci Technol* 37:125–133
40. Shah M, Vakharia V, Chaudhari R, Vora J, Pimenov DY, Giasin K (2022) Tool wear prediction in face milling of stainless steel using singular generative adversarial network and LSTM deep learning models. *Int J Adv Manuf Technol* 121:723–736
41. Cooper C, Wang P, Zhang J, Gao RX, Roney T, Ragai I, Shaffer D (2020) Convolutional neural network-based tool condition monitoring in vertical milling operations using acoustic signals. *Procedia Manuf* 49:105–111
42. Ledig C, Theis L, Huszár F, Caballero J, Cunningham A, Acosta A, Aitken A, Tejani A, Totz J, Wang Z (2017) Photo-realistic single image super-resolution using a generative adversarial network. In: *Proceedings of the IEEE conference on computer vision and pattern recognition (CVPR)*, Honolulu, HI, USA, pp 105–114
43. Cai S, Wang Z, Wang S, Perdikaris P, Karniadakis GE (2021) Physics-informed neural networks for heat transfer problems. *J Heat Transf* 143(6):1–15
44. Cai S, Mao Z, Wang Z, Yin M, Karniadakis GE (2021) Physics-informed neural networks (PINNs) for fluid mechanics: a review. *Acta Mech Sin* 37(12):1727–1738
45. Cuomo S, Di Cola VS, Giampaolo F, Rozza G, Raissi M, Piccialli F (2022) Scientific machine learning through physics-informed neural networks: where we are and what's next. *J Sci Comput* 92(3):88
46. Guo Y, Klink A, Bartolo P, Guo W (2023) Digital twins for electro-physical, chemical, and photonic processes. *CIRP Ann Manuf Technol* 72:593–619
47. Goodfellow I, Pouget-Abadie J, Mirza M, Xu B, Warde-Farley D, Ozair S, Courville A, Bengio Y (2014) Generative adversarial nets. *Adv Neural Inf Process Syst* 27:1–9
48. Zhang S, Guo Y (2009) Taguchi method based process space for optimal surface topography by finish hard milling. *J Manuf Sci Eng* 131(5):051003
49. Quintana G, Ciurana J, Ferrer I, Rodriguez CA (2009) Sound mapping for identification of stability lobe diagrams in milling processes. *Int J Mach Tools Manuf* 49(3–4):203–211
50. Cao H, Yue Y, Chen X, Zhang X (2017) Chatter detection in milling process based on synchrosqueezing transform of sound signals. *Int J Adv Manuf Technol* 89(9):2747–2755
51. De Andrade DC, Leo S, Viana MLDS, Bernkopf C (2018) A neural attention model for speech command recognition. *arXiv preprint. arXiv:1808.08929*
52. Wang Y, Skerry-Ryan R, Stanton D, Wu Y, Weiss RJ, Jaitly N, Yang Z, Xiao Y, Chen Z, Bengio S (2017) Tacotron: towards end-to-end speech synthesis. *INTERSPEECH 2017*, Stockholm, Sweden. *arXiv preprint. arXiv:1703.10135*
53. Zhu J-Y, Park T, Isola P, Efros AA (2017) Unpaired image-to-image translation using cycle-consistent adversarial networks. In: *Proceedings of the IEEE international conference on computer vision (ICCV)*, Venice, Italy, pp 2242–2251
54. Ngiam J, Khosla A, Kim M, Nam J, Lee H, Ng AY (2011) Multimodal deep learning. In: *Proceedings of the 28th international conference on machine learning (ICML-11)*, Bellevue, WA, USA.
55. Phan H, Xie Y, Liao S, Chen J, Yuan B (2020) Cag: a real-time low-cost enhanced-robustness high-transferability content-aware adversarial attack generator. In: *Proceedings of the AAAI conference on artificial intelligence*, New York, New York 34(04):5412–5419

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.