Journal of Manufacturing Systems xxx (xxxx) xxx

Contents lists available at ScienceDirect

Journal of Manufacturing Systems

journal homepage: www.elsevier.com/locate/jmansys

Surface roughness prediction through GAN-synthesized power signal as a process signature☆

Clayton Cooper [a], Jianjing Zhang [a], Y.B. Guo [b,c], Robert X. Gao [a,*]

[a] Department of Mechanical and Aerospace Engineering, Case Western Reserve University, Cleveland, OH, USA
[b] Department of Mechanical and Aerospace Engineering, Rutgers University-New Brunswick, Piscataway, NJ, USA
[c] New Jersey Advanced Manufacturing Institute, Rutgers University-New Brunswick, Piscataway, NJ, USA

ARTICLE INFO

Keywords:
Machining
Surface roughness
Process signature
Generative adversarial network
Convolutional neural network
Machining power

ABSTRACT

Predicting machined surface roughness is critical for estimating a part's performance characteristics such as susceptibility to fatigue and corrosion. Prior studies have indicated that power consumed at the tool-chip interface may represent an indicator for the surface integrity of the machining process. However, no quantitative association has been reported between the machining power and surface roughness due to a lack of data to develop predictive models. This paper presents a data synthesis method to address this gap. Specifically, a conditional generative adversarial network (CGAN) is developed to synthesize power signals associated with varying process parameter combinations. The quality of the synthesized signals is evaluated against experimentally measured power signals by examining the consistency in: 1) the spatial pattern of the signals induced by the cutting process as shown in the frequency domain, and 2) the temporal pattern as shown in the clustering of the synthesized and measured signals corresponding to the same parameter combination. The synthesized signals are then used to augment the measured signals and develop a convolutional neural network (CNN) for predicting the machined surface roughness. Experiments performed using H13 tool steel have shown that data augmentation by CGAN has effectively reduced the error of the surface roughness prediction from 58 %, when no synthetic data is used for CNN training, to 9.1 % when 250 synthetic samples are used. The results demonstrate the effectiveness of CGAN as a data augmentation method and CNN for mapping machining power to surface roughness.

1. Introduction

Surface roughness is one of the most important quality factors to be quantified in machining because it determines part functionality in terms of fatigue resistance, corrosion susceptibility, and fluidic drag [1–3]. Traditional roughness quantification relies on post-process inspection after the workpiece is fabricated by means of contact and contactless profilometry, which often constitutes a bottleneck in manufacturing in terms of time and cost [4,5]. To eliminate this bottleneck, in-process approaches that associate surface roughness with process parameters have been investigated.

In recent years, statistically significant influences of process parameters such as speed, feed, and depth-of-cut (DOC) on surface roughness have been studied [6–8]. However, many of the theoretical approaches do not reflect experimentally observed behavior. For instance, a roughness prediction equation is developed in [9] that relates roughness to feed and tool nose radius, which does not include other machining parameters, material properties, or stochastic process disturbances. This contradicts experimental results in [10] which show that surface roughness depends on the specific cutting energy, which is a function of all process parameters as well as the workpiece condition.

To comprehensively consider process parameters as well as in-process data, machine learning (ML) models have been established to predict surface roughness as a function of parameters and sensor data [11,12]. A limitation in parameter-based roughness quantification has been that it does not consider the in-process variations caused by process dynamics such as workpiece material imperfections and changing tool conditions that are commonly seen in real-world manufacturing environments. To overcome this limitation, the incorporation of process signatures into surface roughness prediction models has attracted

☆ 51st SME North American Manufacturing Research Conference (NAMRC 51, 2023)

* Corresponding author.
E-mail address: robert.gao@case.edu (R.X. Gao).

https://doi.org/10.1016/j.jmsy.2023.05.016

significant interest [13].

Acquired during the machining process, process signatures refer to the measured signals that quantify energy conversion as the workpiece responds to applied loads [14,15]. As the energy-driven part modifications imparted by a machine reflect the compound effect of process parameters and process dynamics, e.g., tool condition and material imperfections that are not accounted for by parameter-based models, process signatures can be used as complementary input to data-driven predictive models [10,16].

Various process signatures, such as sound, force, and temperature, have been investigated. Griffin et al. [17] investigated acoustic emissions as input to a classification tree to predict surface roughness in micromachining. Wu and Lei [18] used features from vibration signals as input to a neural network to predict the roughness of conventionally machined parts. Similarly, Kong et al. [19] used vibration signal features as input to a Bayesian linear regression model for roughness prediction. Wang et al. [20] developed a joint convolutional and recurrent deep learning model to fuse spindle power signals with machined surface images for simultaneous tool wear inference and surface roughness prediction. Tian et al. [21] combined and built upon these approaches to develop a surface prediction model with acoustic, vibrational, force, and power signal features as input to a neural network. A common limitation in the process signatures investigated in these studies is that they only considered a portion of the energy transferred into the workpiece, e.g., vibrational and acoustic dissipations, thus represent an incomplete characterization of the total energy input to the workpiece and its effects on the surface formation. Others, like Moliner-Heredia et al. [22], Wang et al. [23], and Corne et al. [24], have leveraged process signatures to predict tool condition but not the corresponding surface roughness. To establish a direct link between the energy as a process signature and surface roughness and fill an existing research gap, this study presents a data-driven surface roughness predictive model based on energy input only.

Given that specific power measurement instruments are needed to specifically quantify the amount of energy used for material removal during machining, as well as the time-consuming nature of experimental data acquisition, only a limited amount of power/energy data is typically available. To address this limitation, a data synthesis approach has been investigated based on the generative adversarial network (GAN) architecture. GANs rely on a data generator and a discriminator as a pair of competing neural networks for the training process to arrive at an equilibrium point for high-fidelity data synthesis. However, standard GANs operate on random input with no consideration of application-relevant data such as process parameters in machining operations. For this reason, standard GANs do not learn the effects of these variables on the expected output and the synthesized data is difficult to control and interpret. To overcome this challenge, a conditional generative adversarial network (CGAN) has been developed to synthesize power signals based on a limited number of measured power data. In the presented study, the CGAN is regularized by the pertinent machining process parameters to capture the physical effects of varying parameters on the power signal patterns and ensure the output of is physically sound and meaningful.

The validity of the synthesized power signals is then verified from both the physical and the statistical aspects. The former involves quantifying the consistency in the spatial pattern induced by the specific cutting parameters as reflected in the frequency domain of the power signals. The latter reveals the statistical distances among the synthesized intra- and inter-class samples (e.g., samples corresponding to the same or different parameter combinations, respectively), to confirm the capability of CGAN in generating data that exhibit clustering of intra-class samples and separation among inter-class ones.

Finally, in consideration of previous work relating power signal frequencies with tool condition and surface roughness [25], the synthesized signals are decomposed into time-frequency images via wavelet transform and used to train a convolutional neural network (CNN) to predict the roughness of the machined surfaces. A CNN is chosen as the prediction model given its strong image processing capabilities [26] and previous successes analyzing time-frequency data in manufacturing [27]. An overview of the proposed methodology is shown in Fig. 1.

The contributions of this study are summarized as follows:

1. Established a data-driven mapping from the machining energy process signature to surface roughness using CNN, requiring only in-process data as input, and eliminating the need for post-process quantification.
2. Developed a CGAN-based method for high-fidelity synthesis of power signals, which are validated through consistency in signal patterns from both the physical and statistical aspects, thereby alleviating the limitations in data collection for surface predictive modeling.
3. Achieved good surface roughness prediction accuracy by CNN as demonstrated by comparing the synthesized data with experimentally acquired ground truth.

The remainder of the paper is organized as follows: Section 2 provides background on surface roughness and CGANs. Section 3 describes the experiment to evaluate the developed data synthesis and prediction methods. In Section 4, power signal synthesis and surface roughness results are presented and discussed. Conclusions and future research directions are summarized in Section 5.

2. Background and methodology

During machining, material is removed from the workpiece through shear deformation as shown in Fig. 2(a) [28]. Energy is delivered to the workpiece via the motion of the cutting tool that removes a small "chip" of the material. The gross power demand of machining can be modeled as:

$$P_{cut} = MRR(PP)\cdot S_{FC}(PP, TC) + \xi(t, TC, MP) \tag{1}$$

where MRR is the material removal rate (mm^3/s) as a function of related process parameters (PP), S_{FC} is the specific cutting force (N/mm^2) as a function of PP and tool condition (TC), and ξ represents power fluctuations caused by regenerative vibration, etc., which is dependent on time (t), TC, and material properties (MP) [25]. For milling operations, MRR and S_{FC} are defined as:

$$MRR = a_p a_e v \tag{2}$$

$$S_{FC} = k a_p \left[\frac{f_z(1-\cos\phi_s)}{\phi_s}\right]^{\beta+1} \tag{3}$$

where a_p is the axial depth of cut (mm), a_e is the radial depth of cut (mm), v is the cutting feed (mm/s), f_z is the feed per tooth (mm/tooth), ϕ_s is the swept angle of removed material (rad), and k and β are experimentally-derived constants dependent on the tool geometry [29].

Three plots are shown at the bottom of Fig. 2 to illustrate variations in the machining power signal as cutting parameters change: feed in (b), speed in (c), and DOC in (d). In each plot, the solid lines represent measured power signals while the dotted lines represent the theoretical power signals. As seen in these plots, changes in the cutting parameters shift the power curves (both measured and theoretical) up and down, indicating the parameters have a pronounced effect on the power demand of machining. Deviations of the measured power from the theoretical level can be attributed to process variations, i.e., material imperfections, which induce variations through grain boundaries and inhomogeneities that, in turn, cause the cutting forces and corresponding power signal to change. As these imperfections are located randomly throughout the material being machined, their manifestation in the power signals (as deviations from theoretical levels) also exhibit randomness.

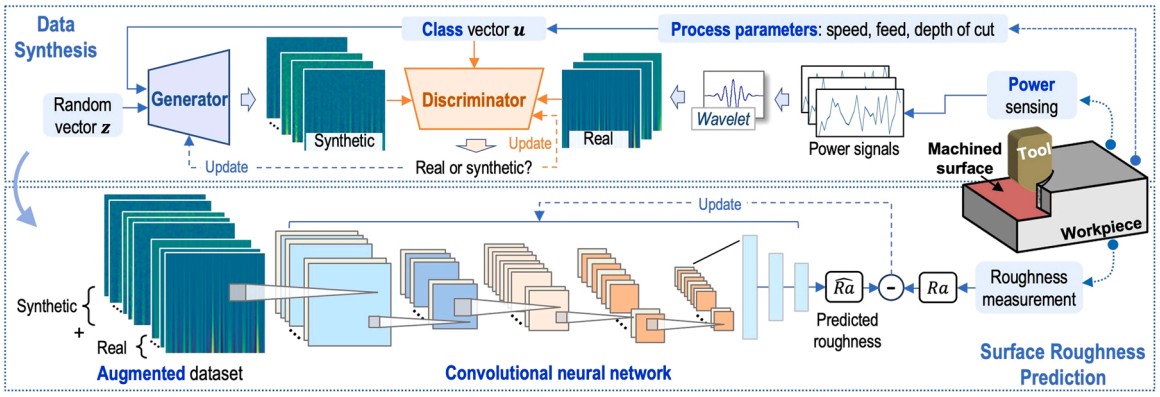

Fig. 1. Surface roughness prediction based on augmented power signals and machine learning.

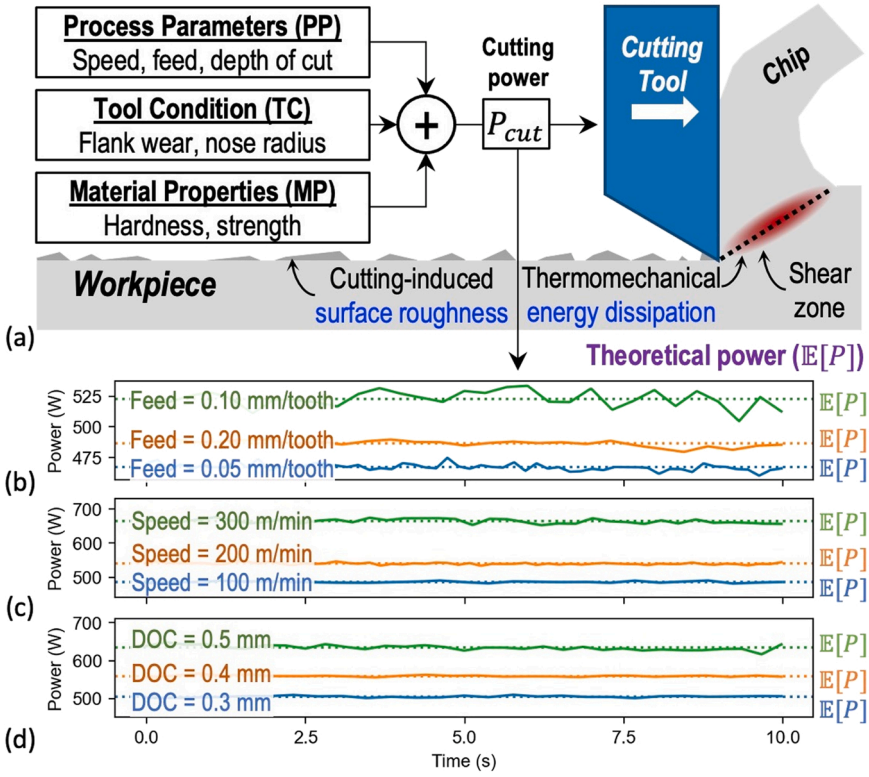

Fig. 2. Chip removal and surface roughness formation during cutting (a); examples of cutting power demand fluctuations over breadth of cutting parameters (b–d); DOC: radial depth-of-cut.

During the material removal process, variations in the material as described above as well as dulling along the cutting edge cause plowing or "ripping" instead of shearing at the microscopic level [30]. This tensile mode of material removal induces a rough surface profile behind the cutting tool as shown in Fig. 3. This surface formation is inherent to all mechanical material removal processes, including milling, turning, and grinding, and indicates the importance of roughness prediction across a breadth of metal removal processes.

The surface roughness quantity of interest in this study is the mean deviation from the profile centerline, *Ra*:

$$Ra = \frac{1}{l_1 - l_0}\int_{l_0}^{l_1} |h(l) - \bar{h}|dl \tag{4}$$

where l_0 and l_1 are the starting and stopping point of the stylus used to trace the surface, respectively, h is the traced surface profile, and $\bar{h}$ is the profile's mean value [31]. Because surface roughness is determined by

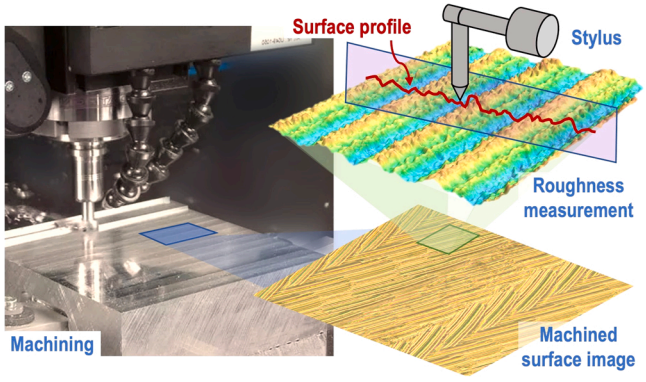

Fig. 3. Machined surface and roughness measurement using stylus (contact profilometry).

the energy transfer at the tool-chip interface, which is in turn determined by the machining power, it is hypothesized that machining power can be mapped to surface roughness by way of a data-driven model using ML methods to allow for in-process roughness inference instead of time-consuming post-process inspection. However, to develop a proper ML model, an adequate power signal dataset covering a broad range of cutting parameters under various operational conditions is needed, which is time-consuming to collect experimentally [32]. To alleviate this limitation and support energy-based data-driven predictive modeling, the development of CGAN for high-fidelity power data synthesis constitutes the focus of this study.

2.1. Fundamentals of GAN

The objective of generative modeling is to approximate the statistical distribution of real-world data (e.g., power signals) using a surrogate model. This distribution can then be sampled to generate new observations, thereby alleviating data availability constraints. GANs have attracted significant attention for this purpose due to their ability to leverage ML to closely learn relevant data distribution and produce realistic synthetic output [33].

The basic structure of a GAN is shown in Fig. 4. Generator G receives random vector z and outputs synthetic data $\tilde{y}$. Meanwhile, discriminator D randomly receives $\tilde{y}$ or ground truth data y and distinguishes the input as either synthetic or real, encoded as 0 and 1, respectively. During the training process, the discriminator's goal is to minimize the classification error whereas the generator's goal is to maximize the classification error by outputting realistic images that mislead the discriminator, i.e., G attempts to learn a surrogate to the unknown ground truth data distribution Ω_y. As a result, G and D are adversaries to one another and GAN training is a minimax game represented by the optimization problem as expressed in (5):

Specifically, the first term of the sum in (5) measures how well the discriminator performs on the ground truth samples. In this scenario, the output from the discriminator contributes to the loss function only when it classifies a real sample as synthetic, or $D(\boldsymbol{y}) = 0$. This is achieved through the logarithm of the discriminator's output such that $\log(D(\boldsymbol{y}))$ $= -c$ (contribution of -c to the loss function), where c is a constant by convention. Similarly, the second term measures how well the generator output can mislead the discriminator. In this case, the output from the discriminator contributes to the loss function only when it classifies a synthetic image as real, or $\log(1 - D(G(z))) = -c$.

By minimizing the second term through training the generator, the network weights are adjusted in the direction of reducing $\log(1 - D(G(z)))$, thus making the generator produce synthetic samples that are less likely to be detected by the discriminator. In contrast, by maximizing both the first and the second terms through training the discriminator, the network weights are adjusted in the direction of increased $\log(D(\boldsymbol{y}))$ and $\log(1 - D(G(z)))$, making both closer to 0 (corresponding to a perfect discriminator performance). Such adversarial training is envisioned to arrive at a Nash or proximal equilibrium where the discriminator can no longer distinguish the synthetic data from the real ones, and the generator can synthesize high-fidelity data but can no longer improve either, due to the lack of further guidance from the discriminator [33,34]. In practice, optimization (5) is carried out in an average sense, by using expectations rather than individual samples.

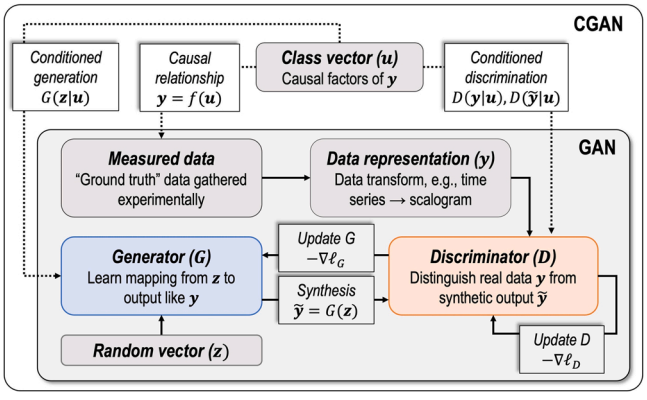

Fig. 4. Fundamental structure of GAN and CGAN.

$$G^{'},\ D^{'} = \min_G \max_D \overbrace{\mathbb{E}_y[\log D(\boldsymbol{y})] + \underbrace{\mathbb{E}_z[\log(1 - D(G(z)))]}_{\ell_G}}^{\ell_D} \tag{5}$$

where G' and D' are optimized G and D, and the denoted loss quantities (ℓ_*) are the objective functions when training the generator and discriminator models.

In recent years, GANs have emerged as effective tools for data synthesis in manufacturing and related fields [35]. For example, Hertlein et al. [36] developed a 2-D CNN GAN for topology optimization in additive manufacturing to improve manufacturability, which was shown to be significantly faster than traditional topology optimization methods. Likewise, Schleich et al. [37] formulated a 2-D CNN GAN for tolerance analysis that predicted shape-agnostic part deviations and enabled expected error stack-up to be calculated before production took place. Cooper et al. [38] used a 2-D CNN GAN to synthesize milling audio signals as spectrograms and develop a tool condition classification model using a mix of experimental and synthetic data. Barua et al. [39] proposed a joint FC and 2-D architecture for multi-view object synthesis that was shown to be extensible over a breadth of datasets and learning problems. Qu et al. [40] and Liu et al. [41] developed GANs to synthesize speed, temperature, and vibration data for wind turbines using simultaneous, cross-sensor 1-D convolution operations such that a sensor fusion fault detection system could be developed. By generating outputs based on limited training data, GANs have shown to reduce the need for extensive experimentation, conserve resources, and improve the robustness of predictive models by expanding the data availability for developing such models.

2.2. CGAN for synthesis of cutting power profile

Numerous variants to the standard GAN structure have been proposed over recent years for a breadth of learning tasks. Of these, CGANs are of significant interest to this study [42]. A prevalent issue with the standard GAN structure is that the generator output cannot be conditioned on the physical application-specific classes or labels. Since varying process parameters have a direct impact on the formation of surface roughness as shown in Fig. 2, the synthesized power signals are required to contain parameter-specific patterns to retain the roughness-related information embedded in different process parameter combinations (e.g., classes) to ensure the validity of predictive modeling.

To solve this problem and enable class-dependent data synthesis, CGAN is investigated in which the generator and discriminator outputs are conditioned on additional class information $\boldsymbol{u}$. Consequently, Eq. (5) is modified as

$$G^{'},\ D^{'} = \min_G \max_D \overbrace{\mathbb{E}_y[\log D(\boldsymbol{y}|\boldsymbol{u})] + \underbrace{\mathbb{E}_z[\log(1 - D(G(z|\boldsymbol{u})|\boldsymbol{u}))]}_{\ell_G}}^{\ell_D} \tag{6}$$

where conditioning is achieved by passing $\boldsymbol{u}$ as an input to both the generator and the discriminator in addition to z or $\boldsymbol{y}$, as shown in Fig. 4 This enables the generator to learn a unique output distribution corresponding to each $\boldsymbol{u}$ while also allowing the discriminator to conduct class-specific classification. In other words, G learns surrogates of $\Omega_1(\boldsymbol{u}_1)$ $,\Omega_2(\boldsymbol{u}_2),\ldots,\Omega_N(\boldsymbol{u}_N)$, where $\Omega_n(\boldsymbol{u}_n)$ is the ground truth data distribution of the n^{th} class and N is the number of classes used to train the CGAN. D then learns to determine whether the synthesized sample

$G(z|\boldsymbol{u}_n)$ belongs to $\Omega_n(\boldsymbol{u}_n)$.

Training of a CGAN structure follows the same procedure as that of a conventional GAN. Once trained, the generator can be queried with $(z, \boldsymbol{u})$ pairs to synthesize power signals corresponding to the class input $\boldsymbol{u}$ and augment the dataset. For this study, $\boldsymbol{u}$ is a vector representing process speed, feed, and radial depth-of-cut, since these physical attributes have been shown to effectively characterize the power signal [43].

3. Experimental evaluation

To validate the proposed surface roughness prediction method using CGAN data augmentation, milling experiments have been performed at various speeds, feeds, and depths-of-cut to collect a set of spindle power signals.

3.1. Milling data collection and representation

A total of 21 prismatic specimens of AISI H13 tool steel (50 ± 1 HRC) were dry milled on a CNC machining center using a 20 mm, two flute end mill with tungsten carbide inserts. Six cuts were made using each set of the cutting parameters listed in Table 1, yielding 126 experimental data. Spindle power was recorded during each cut at 341 kHz using a Fluke Norma power analyzer. Following milling, the average *Ra* value in the stepover direction is determined for each combination of cutting parameters using contact profilometry.

Previous studies [25] have shown that low-frequency components of machining power signals dissipate as tool wear increases. Since tool wear and surface roughness are closely linked, it is expected that the frequency spectrum of the power signals contains surface roughness-related information. Thus, to ensure the CGAN generates realistic power data in both the time and frequency domains of the signal, continuous wavelet transform is first used to represent the power signals as time-frequency images, or scalograms [44]. The Ricker wavelet was selected as the base wavelet, based on the criterion of best energy-to-entropy ratio as described in [45].

Representative wavelet scalograms for specimens #1–9 are shown in Fig. 5 with the corresponding power signals. In each scalogram, the peaks and valleys present along the time axis correspond to sudden changes in the power demand caused by vibration, material imperfection, and tool deterioration. These events manifest themselves in the time-frequency domain as slender spikes reaching into the upper frequency bands but quickly disappearing as the transient event ends. It is noted that the occurrence of these transient events forms different distributions under different parameter combinations. For example, the occurrence exhibits a more regular pattern for specimens #1 and #8, while appearing more sparsely in specimens #2, 5, 6, and 7. The distinct patterns in distributions indicate the underlying influence of process parameters when coupled with process variations, and they can be verified physically against the ground truth by comparing the measured and synthetic signals in the frequency domain, as well as statistically, by assessing the statistical similarity between the real and synthetic data using embedding techniques.

Table 1
Milling parameters and measured surface roughness.

Spec.	Speed (mm/min)	Feed/tooth (mm/tooth)	Radial DOC (mm)	*Ra* (nm)
1	200	0.1	0.3	241
2	200	0.1	0.5	520
3	200	0.1	0.5	384
4	200	0.05	0.5	360
5	200	0.1	0.4	646
6	100	0.1	0.5	440
7	200	0.1	0.5	480
8	300	0.1	0.5	650
9	200	0.05	0.5	372
10	200	0.1	0.4	329
11	100	0.1	0.5	399
12	300	0.1	0.5	396
13	200	0.1	0.3	426
14	200	0.05	0.5	374
15	200	0.2	0.5	488
16	100	0.1	0.5	356
17	200	0.1	0.3	265
18	200	0.1	0.4	399
19	200	0.2	0.3	584
20	300	0.1	0.5	657
21	200	0.2	0.5	1059

3.2. Quantification of generated image quality

To quantify consistency of the synthesized image signals with process physics and statistical faithfulness to the ground truth data, two metrics are considered: 1) the frequency spectra of the generated signals as compared to the ground truth, and 2) inter- and intra-class similarity as quantified by the method of uniform manifold approximation and projection (UMAP).

It is known from previous work that process parameter settings uniquely determine the frequency content of the machining power signals [25]. For instance, the dominant frequency in the power signal monotonically increases as cutting speed increases. The same relationship holds for the feed/tooth. Each set of parameters ($\boldsymbol{u}$) should therefore have a unique frequency spectrum associated with it. This observation serves as a basis of comparison for experimental and synthetic power signals, namely the frequency spectra of the real and synthesized data should be similar under the same process parameters. This similarity can be quantified as:

$$\xi_{freq} = \rho\left(\mathscr{F}(\boldsymbol{P}), \mathscr{F}\left(\widetilde{\boldsymbol{P}}\right)\right) \tag{7}$$

where ξ_{freq} is the spectral similarity metric, ρ is the Pearson correlation [46] of the Fourier spectra $\mathscr{F}(\boldsymbol{P})$ and $\mathscr{F}\left(\widetilde{\boldsymbol{P}}\right)$, $\boldsymbol{P}$ is a ground truth power signal, $\widetilde{\boldsymbol{P}}$ is a synthetic signal found via inverse wavelet transform [44] of the generated scalogram, and both signals share the same $\boldsymbol{u}$.

If the generator has correctly established the unique mapping between the frequency distribution of the power signal and the physical information embedded in the process parameters, correlation between the real and synthetic frequency spectra will approach 1. This metric, however, speaks only to each datum and does not assess if unique $\Omega_n(\boldsymbol{u}_n)$ were learned for each combination of process parameters. For this purpose, UMAP is investigated.

The fundamental concept of UMAP is to visualize the distribution of high-dimensional data, e.g., scalogram images (or time-frequency decomposition of the power signals) in this study, by learning a low-dimensional surrogate distribution which maintains the distance between the data points proportional to their high-dimensional counterparts. As a result, statistically similar data clusters together whereas dissimilar data separate. Details of UMAP are provided in [47].

To verify the result of power signal synthesis, the synthetic scalograms conditioned on the same process parameter combination should cluster together while maintaining separability from those conditioned on different parameter combinations. Since different CGAN architectures will yield different spectral correlations and UMAP embeddings, several CGAN architectures are investigated.

3.3. CGAN architecture search

To examine the generator behavior across different CGAN architectures, experiments have been performed to train fully connected (FC), 1-D CNN, and 2-D CNN CGANs. A hyperparameter search is performed for each model independently over the ranges listed in Table 2 to determine the optimal network structure and training process. These same ranges are used to optimize the CNN using backpropagation once the generator

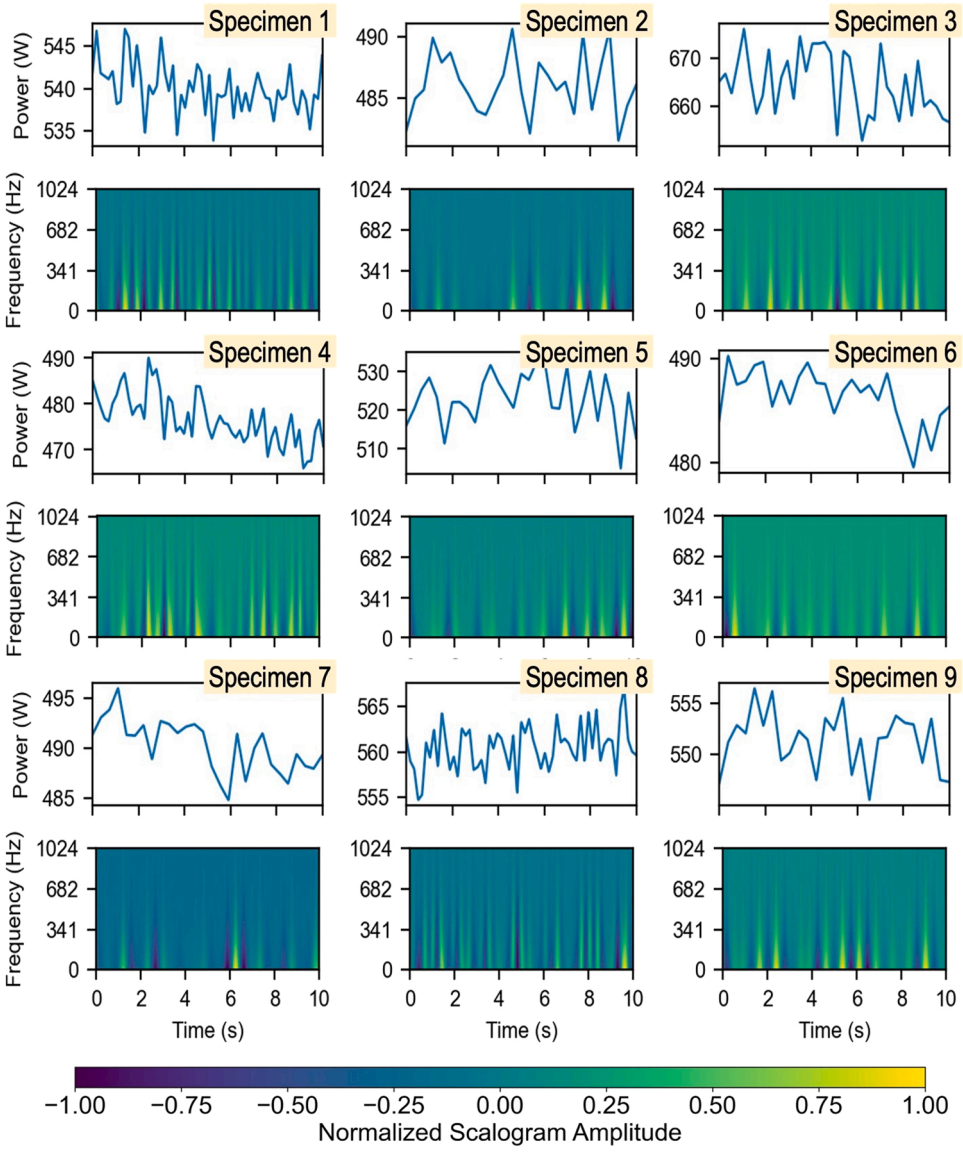

Fig. 5. Experimentally measured power signals and scalograms.

Table 2
ML model hyperparameters and search ranges; SGD: stochastic gradient descent, PReLU: parametric rectified linear unit.

Hyperparameter	Range
Optimizer	{Adam, RMSprop, SGD}
Initial learning rate	$\{1 \times 10^{-6}, 1 \times 10^{-5}, \ldots, 1 \times 10^{-2}\}$
Activation function	{Leaky ReLU, PReLU}
Batch size	$\{2^0, 2^2, \ldots, 2^6\}$
Dropout probability	{0 %, 20 %, …, 80 %}
z length	{200, 400, …, 1000}
Number of layers	{2, 3, …, 6}
Kernel size (CNN)	{3, 5, …, 11}
Kernels/layer (CNN)	$\{2^5, 2^6, \ldots, 2^9\}$
Layer size (FC)	$\{2^{10}, 2^{11}, \ldots, 2^{14}\}$

and discriminator are trained, also using backpropagation [48]. The generator and the discriminator are constrained to the same architecture category during the search.

3.4. Surface roughness prediction

To predict surface roughness, each CGAN is trained to its equilibrium and its generator is used to generate k images for each specimen, 70 % of which are used to train a CNN for surface roughness prediction, and the remaining 30 % are held out to validate the model after training. (The effects of varying k from 0 to 500 images/specimen are shown for the CNN and discussed in Section 4.) A CNN is chosen as the surface roughness prediction model due to its ability to extract features from the scalograms using sequential convolutions and fuse the features from low-level to high-level to yield a surface roughness estimate $\widehat{Ra}$ [49].

Surface roughness prediction error is quantified as the mean absolute percentage error (MAPE), which is iteratively minimized for a training dataset using gradient descent and assessed after training on a validation dataset that is held out of the training data [50].

4. Results and discussion

4.1. CGAN-synthesized power signal

Exemplary data synthesis results associated with each CGAN are

shown in Fig. 6. While the synthesized scalograms from both the FC-CGAN and the 1-D CNN CGAN exhibit a slender spike pattern similar to the ground truth images, they are much more defined for the FC model than the 1-D CNN with substantially less noise contamination. Notably, while the FC model retains qualitative similarity, e.g., spike shape and color, the spikes are shown to translate left and right along the time axis as compared to the ground truth, indicating that the GAN is synthesizing transient events in the power signal as seen in Fig. 5, but at variable times. Meanwhile, the 2-D CNN failed to converge for any hyperparameter combination as described by the search ranges (Table 2). Oscillatory behavior in the 2-D CNN generator and discriminator loss functions was observed for learning rates between 0.000001 and 0.001. As a result, no meaningful patterns were recognized from the generated scalograms.

The performance of the CGANs is quantitatively evaluated first using frequency spectra correlation between the synthesized power signal (obtained using inverse wavelet transform of scalogram from each CGAN model) and the ground truth, measured signal for each specimen, as depicted in Fig. 7. For the FC CGAN, the average frequency spectra correlation is greater than 0.9 across all specimens, indicating that the frequencies present in the synthetic signals closely resemble that of the ground truth. The results from 1-D CNN CGAN yield correlations in the range of 0.6–0.8. This is consistent with the results shown in Fig. 5 that the scalogram appearing in the synthetic 1-D CNN images still captures the frequency patterns associated with different process parameter combinations but is obscured by noise. Since the 2-D CNN images do not converge to any frequency pattern, their spectral correlation coefficients are near zero as expected.

The ineffectiveness of the 2-D CNN can be considered as caused by the poor performance of the convolutional generators due in part to the sparse connections within the network. Whereas the FC generator considers every pixel when passing the input through the network layers, the convolutional models only consider local regions of 3–11 pixels at once. Thus, the FC network can guide training for the entire image structure at once whereas the CNNs analyze only local patches. Additionally, the number of learnable parameters is higher for the FC model than the 1-D and 2-D CNNs (178 M vs. 1.1 M and 2.8 M, respectively), indicating stronger learning potential. Further investigation of the learning capability of different CGANs is considered for future studies.

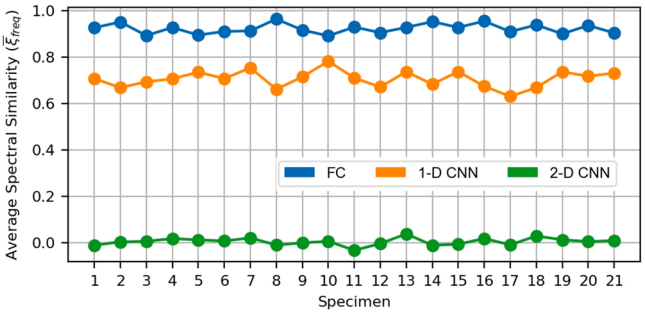

Fig. 7. Frequency spectra correlation with ground truth for different CGAN architectures.

In addition to the established consistency between the synthesized data and the ground truth in terms of frequency patterns, the capability of the developed CGAN in learning parameter-specific distributions $\Omega_n(\boldsymbol{u}_n), n = 1, 2, \ldots 21$ is also evaluated using the UMAP embedding.

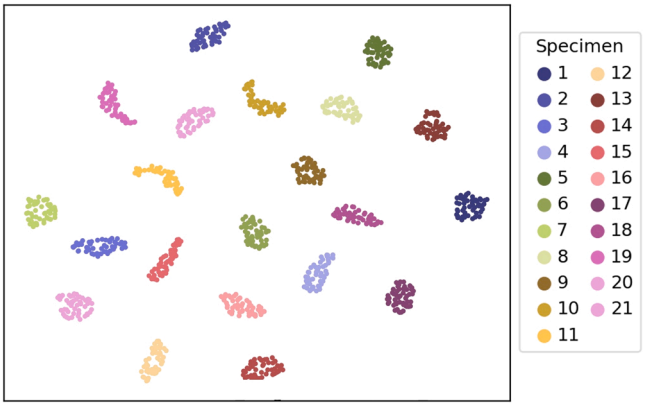

Fig. 8. UMAP embeddings of FC CGAN images (100/specimen) showing clustering of intra-class samples and separation of inter-class ones; UMAP axes are dimensionless.

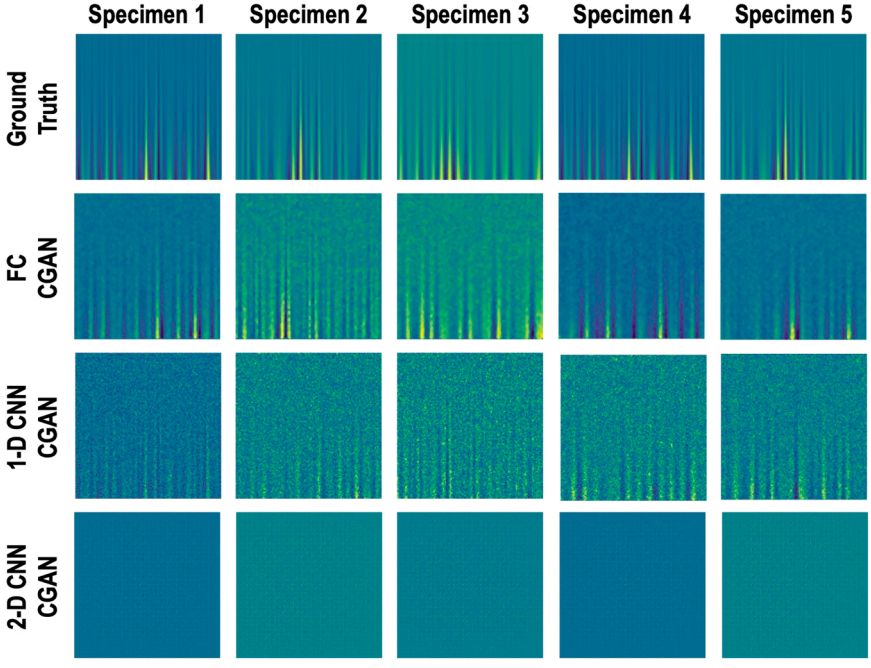

Fig. 6. Synthesized scalograms from different CGAN architectures.

Fig. 8 shows the distribution of 2100 synthesized images corresponding to 21 different process combinations in the UMAP embedding space. Each sample is color-coded according to the class ***u*** at the generator's input. It is seen that intra-class clustering and inter-class separation have been achieved, confirming a well-trained generator network that is able to generate images unique to each process parameter combination. The structure of the best-performing FC CGAN is shown in Table 3. The optimizer during training is Adam with a learning rate of 0.01. The activation function is Leaky ReLU.

4.2. Surface roughness prediction results

The optimized CNN structure for surface roughness prediction is tabulated in Table 4, trained with Adam optimizer (learning rate 0.01) and Leaky ReLU activation function. The prediction error (in MAPE) is shown in Fig. 9 as a function of the size of the synthetic data. During this experiment, the training dataset size was the only variable while the CNN architecture and hyperparameters were held constant. With all other variables fixed, a larger dataset size produced a lower prediction error because more diverse synthetic training data enabled the CNN to fully optimize the network weights and learn the unique patterns from the scalograms. It is seen that the surface roughness prediction error is reduced from 58 %, when only the ground truth data but no synthetic data was used to train the CNN, to 9.1 % when 250 or more synthetic samples are included for CNN training.

In addition to the synthetic data size, the effect of the ground truth data size is also evaluated on power signal synthesis and surface roughness prediction. Specifically, the CGAN is retrained using progressively smaller ground truth datasets, ranging from 6 images/specimen to 2 images/specimen, once again holding constant the CGAN architecture and hyperparameters. The resulting MAPE of the roughness prediction are shown in Fig. 10. It is seen that increasing the number of ground truth samples per specimen significantly improves CGAN image sharpness as well as the accuracy of the surface roughness prediction. For instance, images synthesized using 3 ground truth samples for CGAN training are significantly blurrier and more amorphous than the scenario with 6 ground truth training samples per specimen, which produces sharper images with more detail in the higher frequency bands. Additionally, MAPE is reduced by 56 % when the ground truth dataset is increased from 3 samples/specimen to 6 samples/specimen, indicating that the impact of ground truth dataset size is significant in determining CGAN and CNN behavior in the developed surface roughness prediction method.

The surface roughness prediction method developed in this work is compared to representative work by other researchers on surface roughness prediction, and the result is summarized in Table 5. With the exception of [21], this study's technique exhibits a lower prediction error than other comparable studies. However, the CGAN + CNN approach relies on only a single process signature as input rather than the four upon which [21] relies, making the former approach more feasible than the latter in a production setting. Additionally, the BGWO model requires extraction of 73 statistical features from the four input signatures, which is computationally more complex than the CGAN + CNN approach as presented in this study. Nonetheless, low prediction error of [21] indicates the effect of multi-sensor fusion, which will be considered as one of the authors' research topics in the future.

Table 3
Optimal network structure of FC CGAN.

	Generator	Discriminator
Input	**z**: 1000, **u**: 200	Scalogram: 16,384, **u**: 200
Layer #1	2048 (2^{11})	2048 (2^{11})
Layer #2	4096 (2^{12})	1024 (2^{10})
Layer #3	8192 (2^{13})	512 (2^{9})
Output	16,384 (128^{2})	1
Dropout %	60 %	40 %
Batch Size	64	

Table 4
Optimal network structure of 2D CNN.

Input Size	128 × 128
Kernel #1	11 × 11×64
Kernel #2	7 × 7×128
Kernel #3	5 × 5×256
Kernel #4 and #5	3 × 3×512
FC #1	4608
FC #2	1024
FC #3	64
Output	1
Batch Size	64

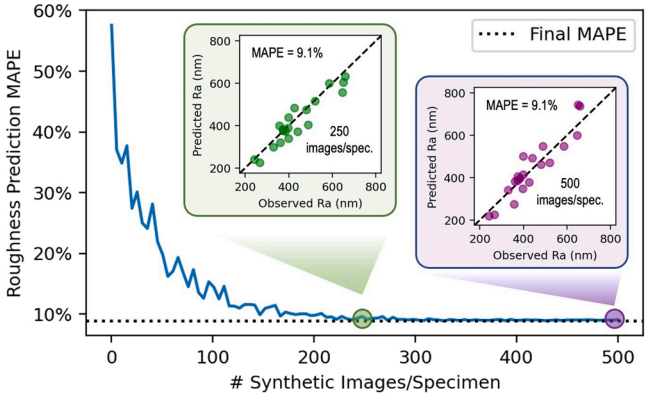

Fig. 9. Roughness prediction error vs. synthetic data size; insets depict measured vs. CNN-predicted *Ra* where dashed line is ideal.

5. Conclusion

In an effort to establish a direct mapping between power signals and surface roughness in end milling that enables in-process roughness quantification and eliminates the bottleneck of post-process inspection, a data-driven approach based on CGAN and CNN has been developed. The CGAN enables synthesis of high-fidelity power signals under varying process parameter combinations to circumvent the limitation of power signal data availability and augment the training dataset. The CNN then takes the augmented dataset and learns the surface roughness by way of the power signals' time-frequency characteristics as extracted via wavelet transform. An experiment has been carried out on milling H13 tool steel to evaluate the developed methods.

Examining both the physical and statistical patterns of the synthesized power data using frequency spectra and UMAP, respectively, it is confirmed that CGAN is to learn unique data distributions specific to each process parameter combination. The CNN trained on the augmented dataset has achieved a 9.1 % mean error in predicting surface roughness, confirming the power process signature as a viable candidate for in-process surface characterization in machining. It is further observed that the prediction error is asymptotic as a function of the synthetic dataset size, with a larger ground truth dataset resulting in smaller prediction error.

Industrial application of our method is envisioned to follow a two-step process: 1) model recalibration for the process and machine of interest, and 2) provisioning of real-world, in-situ production data to refine our developed model. Regarding point 1, the CGAN and CNN from our method should be recalibrated to the application-specific power signals and process parameters to avoid extrapolating beyond the range of the training data, which may yield spurious results. The model recalibration process to avoid these issues would involve collecting a small number of power signals and surface roughness measurements,

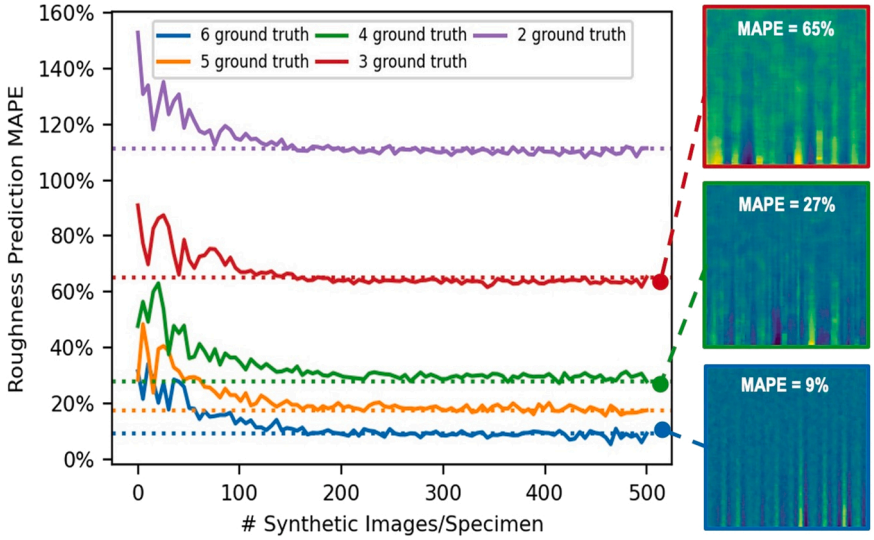

Fig. 10. Surface roughness prediction error under reducing CGAN training dataset sizes.

Table 5
Comparison of presented study to literature; BP ANN: backpropagation artificial neural network, BLR: Bayesian linear regression, RNN: recurrent neural network, BGWO: binary grey wolf optimization.

Ref.	Process Signature Input	R_a Prediction Model	R_a Prediction Error
Ours	Power profile	CGAN + CNN	9 %
[18]	Spindle, fixture vibrations	BP ANN	18 %
[19]	Spindle, fixture, workpiece vibrations	BLR	10 %
[20]	Surface image, tool image, power profile	CNN + RNN	12 %
[21]	Acoustic, spindle vibration, fixture forces, power profile	BGWO	1 %

such as the 126 observed in our study, and retraining the CGAN and CNN using the procedure described in Section 3.3. Regarding point 2, the post-calibration CNN from point 1 will be used for surface roughness prediction during production by providing it with time-frequency images of in-situ power signals acquired by the end user. This implementation scheme represents a low barrier to entry for industrial use of the developed techniques as presented in this study.

Future research will explore the potential of generative models in synthesizing additional process signatures. Additionally, tool wear will be included in the CGAN input to more comprehensively account for process variations in surface characterization. The effects of changing tool material and geometry on the scalogram synthesis and surface roughness prediction will also be studied in order to make the developed model more generalizable across tool types. Finally, physical factors of the machining process such as strong plastic deformation and chatter will be detected in the power signals and accounted for by future surface roughness prediction models to further reduce the prediction error.

Acknowledgments

This work is supported by the National Science Foundation under grant CMMI-2040288/2040358. Clayton Cooper acknowledges support from the National Science Foundation Graduate Research Fellowship under Grant No. 1937968. Robert Gao and Jianjing Zhang acknowledge support from the NSF Engineering Research Center for Hybrid Autonomous Manufacturing: Moving from Evolution to Revolution (ERC-HAMMER) under award EEC-2133630.

References

[1] Toloei A, Stoilov V, Northwood D. The relationship between surface roughness and corrosion. In: Volume 2B: Advanced Manufacturing, San Diego, California, USA: American Society of Mechanical Engineers; Nov. 2013, p. V02BT02A054. ⟨https://doi.org/10.1115/IMECE2013-65498⟩.
[2] Gu X, Cegla F. The effect of internal pipe wall roughness on the accuracy of clamp-on ultrasonic flowmeters. IEEE Trans Instrum Meas 2019;68(1):65–72. https://doi.org/10.1109/TIM.2018.2834118.
[3] Zhao B, Song J, Xie L, Hu Z, Chen J. Surface roughness effect on fatigue strength of aluminum alloy using revised stress field intensity approach. Sci Rep 2021;11(1):19279. https://doi.org/10.1038/s41598-021-98858-0.
[4] Quick Guide to Surface Roughness Measurement Mitutoyo; Dec. 2016.
[5] Ghodrati S, Kandi SG, Mohseni M. Nondestructive, fast, and cost-effective image processing method for roughness measurement of randomly rough metallic surfaces. J Opt Soc Am A 2018;35(6):998. https://doi.org/10.1364/JOSAA.35.000998.
[6] Chen C-H, Jeng S-Y, Lin C-J. Prediction and analysis of the surface roughness in CNC end milling using neural networks. Appl Sci 2021;12(1):393. https://doi.org/10.3390/app12010393.
[7] Yeganefar A, Niknam SA, Asadi R. The use of support vector machine, neural network, and regression analysis to predict and optimize surface roughness and cutting forces in milling. Int J Adv Manuf Technol 2019;105(1–4):951–65. https://doi.org/10.1007/s00170-019-04227-7.
[8] Sekulic M, Pejic V, Brezocnik M, Gostimirovic M, Hadzistevic M. Prediction of surface roughness in the ball-end milling process using response surface methodology, genetic algorithms, and grey wolf optimizer algorithm. Adv Prod Eng Manag 2018;13(1):18–30. https://doi.org/10.14743/apem2018.1.270.
[9] Groover MP. Fundamentals of Modern Manufacturing: Materials, Processes, and Systems. 4th ed. Hoboken, NJ: J. Wiley & Sons; 2010.
[10] Sealy MP, Liu ZY, Guo YB, Liu ZQ. Energy based process signature for surface integrity in hard milling. J Mater Process Technol 2016;238:284–9. https://doi.org/10.1016/j.jmatprotec.2016.07.038.
[11] Nasir V, Sassani F. A review on deep learning in machining and tool monitoring: methods, opportunities, and challenges. Int J Adv Manuf Technol 2021;115(9–10):2683–709. https://doi.org/10.1007/s00170-021-07325-7.
[12] Pimenov DY, Bustillo A, Wojciechowski S, Sharma VS, Gupta MK, Kuntoğlu M. Artificial intelligence systems for tool condition monitoring in machining: analysis and critical review. J Intell Manuf 2022. https://doi.org/10.1007/s10845-022-01923-2.
[13] Bhuiyan MSH, Choudhury IA, Dahari M. Monitoring the tool wear, surface roughness and chip formation occurrences using multiple sensors in turning. J Manuf Syst 2014;33(4):476–87. https://doi.org/10.1016/j.jmsy.2014.04.005.
[14] Process Signature, Collaborative Research Center 136. German Research Foundation. [Online]. Available: ⟨https://www.prozesssignaturen.de/en/about-us/glossary/p/process-signature⟩ [Accessed: 19 September 2022].
[15] Process Quantities, Collaborative Research Center 136. German Research Foundation. Accessed: Sep. 19, 2022. [Online]. Available: ⟨https://www.prozesssignaturen.de/en/about-us/glossary/p/process-quantities⟩ [Accessed: 19 September 2022].
[16] Ragai I, Abdalla AS, Abdeltawab H, Qian F, Ma J. Toward smart manufacturing: Analysis and classification of cutting parameters and energy consumption patterns in turning processes. S027861252200067X J Manuf Syst 2022. https://doi.org/10.1016/j.jmsy.2022.04.016.

[17] Griffin JM, et al. Control of deviations and prediction of surface roughness from micro machining of THz waveguides using acoustic emission signals. Mech Syst Signal Process 2017;85:1020–34. https://doi.org/10.1016/j.ymssp.2016.09.016.
[18] Wu TY, Lei KW. Prediction of surface roughness in milling process using vibration signal analysis and artificial neural network. Int J Adv Manuf Technol 2019;102 (1–4):305–14. https://doi.org/10.1007/s00170-018-3176-2.
[19] Kong D, Zhu J, Duan C, Lu L, Chen D. Bayesian linear regression for surface roughness prediction. Mech Syst Signal Process 2020;142:106770. https://doi.org/10.1016/j.ymssp.2020.106770.
[20] Wang P, Liu Z, Gao RX, Guo Y. Heterogeneous data-driven hybrid machine learning for tool condition prognosis. CIRP Ann 2019;68(1):455–8. https://doi.org/10.1016/j.cirp.2019.03.007.
[21] Tian W, et al. Broad learning system based on binary grey wolf optimization for surface roughness prediction in slot milling. IEEE Trans Instrum Meas 2022;71: 1–10. https://doi.org/10.1109/TIM.2022.3144232.
[22] Moliner-Heredia R, Peñarrocha-Alós I, Abellán-Nebot JV. Model-based tool condition prognosis using power consumption and scarce surface roughness measurements. J Manuf Syst 2021;61:311–25. https://doi.org/10.1016/j.jmsy.2021.09.001.
[23] Wang J, Li Y, Zhao R, Gao RX. Physics guided neural network for machining tool wear prediction. J Manuf Syst 2020;57:298–310. https://doi.org/10.1016/j.jmsy.2020.09.005.
[24] Corne R, Nath C, El Mansori M, Kurfess T. Study of spindle power data with neural network for predicting real-time tool wear/breakage during inconel drilling. J Manuf Syst 2017;43:287–95. https://doi.org/10.1016/j.jmsy.2017.01.004.
[25] Wang X, Williams RE, Sealy MP, Rao PK, Guo Y. Stochastic modeling and analysis of spindle power during hard milling with a focus on tool wear. J Manuf Sci Eng 2018;140(11):111011. https://doi.org/10.1115/1.4040728.
[26] Krizhevsky A, Sutskever I, Hinton GE. ImageNet classification with deep convolutional neural networks. Commun ACM 2017;60(6):84–90. https://doi.org/10.1145/3065386.
[27] Jiang X, Scott P, Whitehouse D. Wavelets and their applications for surface metrology. CIRP Ann 2008;57(1):555–8. https://doi.org/10.1016/j.cirp.2008.03.110.
[28] Priyadarshini A, Pal SK, Samantaray AK. Finite element modeling of chip formation in orthogonal machining. In: Statistical and Computational Techniques in Manufacturing, J. P. Davim, Ed., Berlin, Heidelberg: Springer Berlin Heidelberg; 2012, pp. 101–144. ⟨https://doi.org/10.1007/978-3-642-25859-6_3⟩.
[29] Liu Z, Guo Y. A hybrid approach to integrate machine learning and process mechanics for the prediction of specific cutting energy. CIRP Ann 2018;67(1): 57–60. https://doi.org/10.1016/j.cirp.2018.03.015.
[30] Bhokse V, Chinchanikar S, Anerao P, Kulkarni A. Experimental investigations on chip formation and plowing cutting forces during hard turning. Mater Today: Proc 2015;2(4–5):3268–76. https://doi.org/10.1016/j.matpr.2015.07.138.
[31] Oberg E, Jones FD, Horton HL, Ryffel H.H., McCauley CJ, Brengelman L. Machinery's handbook, 31st ed. South Norwalk: Industrial Press, Inc; 2020.
[32] Lu F, Zhou G, Liu Y, Zhang C. Ensemble transfer learning for cutting energy consumption prediction of aviation parts towards green manufacturing. J Clean Prod 2022;331:129920. https://doi.org/10.1016/j.jclepro.2021.129920.
[33] Goodfellow IJ et al., Generative adversarial networks. arXiv; Jun. 10, 2014. [Online]. Available: ⟨http://arxiv.org/abs/1406.2661⟩ [Accessed: 06 September 2022].
[34] Farnia F, Ozdaglar A. Do GANs always have Nash equilibria? In: Proceedings of the 37th international conference on machine learning, in Proceedings of Machine Learning Research, 119. PMLR; 2020, pp. 3029–3039.
[35] Ntavelis E, Kastanis I, Van Gool L, Timofte R. Same same but different: augmentation of tiny industrial datasets using generative adversarial networks. In: Proceedings of the 2020 7th Swiss conference on data science (SDS), Luzern, Switzerland: IEEE; Jun. 2020, pp. 17–22. ⟨https://doi.org/10.1109/SDS49233.2020.00011⟩.
[36] Hertlein N, Buskohl PR, Gillman A, Vemaganti K, Anand S. Generative adversarial network for early-stage design flexibility in topology optimization for additive manufacturing. J Manuf Syst 2021;59:675–85. https://doi.org/10.1016/j.jmsy.2021.04.007.
[37] Schleich B, Qie Y, Wartzack S, Anwer N. Generative adversarial networks for tolerance analysis. CIRP Ann 2022;71(1):133–6. https://doi.org/10.1016/j.cirp.2022.03.021.
[38] Cooper C, Zhang J, Gao RX, Wang P, Ragai I. Anomaly detection in milling tools using acoustic signals and generative adversarial networks. Procedia Manuf 2020: 372–8. https://doi.org/10.1016/j.promfg.2020.05.059.
[39] Barua S, Erfani SM, Bailey J. FCC-GAN: A fully connected and convolutional net architecture for GANs. arXiv, May 27; 2019. [Online]. Available: ⟨http://arxiv.org/abs/1905.02417⟩ [Accessed: 14 October 2022].
[40] Qu F, Liu J, Ma Y, Zang D, Fu M. A novel wind turbine data imputation method with multiple optimizations based on GANs. Mech Syst Signal Process 2020;139: 106610. https://doi.org/10.1016/j.ymssp.2019.106610.
[41] Liu J, Qu F, Hong X, Zhang H. A small-sample wind turbine fault detection method with synthetic fault data using generative adversarial nets. IEEE Trans Ind Inf 2019;15(7):3877–88. https://doi.org/10.1109/TII.2018.2885365.
[42] Mirza M, Osindero S. Conditional Generative Adversarial Nets. arXiv; Nov. 06, 2014. [Online]. Available: ⟨http://arxiv.org/abs/1411.1784⟩ [Accessed: 02 September 2022].
[43] Liu ZY, Guo YB, Sealy MP, Liu ZQ. Energy consumption and process sustainability of hard milling with tool wear progression. J Mater Process Technol 2016;229: 305–12. https://doi.org/10.1016/j.jmatprotec.2015.09.032.
[44] Gao RX, Yan R. Wavelets: theory and applications for manufacturing. Boston, MA: Springer US; 2011. ⟨https://doi.org/10.1007/978-1-4419-1545-0⟩.
[45] Yan R, Gao RX. Base wavelet selection for bearing vibration signal analysis. Int J Wavel Multiresolut Inf Process 2009;07(04):411–26. https://doi.org/10.1142/S0219691309002994.
[46] Everitt B, Skrondal A. The Cambridge dictionary of statistics; 2010. [Online]. Available: ⟨http://www.books24x7.com/marc.asp?bookid=36106⟩ [Accessed: 08 December 2021].
[47] McInnes L, Healy J, Melville J. UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction." Sep. 17; 2020. [Online]. Available: ⟨http://arxiv.org/abs/1802.03426⟩ [Accessed: 11 August 2021].
[48] Rojas R. The backpropagation algorithm. In: Neural Networks, Berlin, Heidelberg: Springer Berlin Heidelberg; 1996, pp. 149–182. ⟨https://doi.org/10.1007/978-3-642-61068-4_7⟩.
[49] LeCun Y, Bengio Y, Hinton G. Deep learning. Nature 2015;521(7553):436–44. https://doi.org/10.1038/nature14539.
[50] Malakooti B. Operations and production systems with multiple objectives. Hoboken, New Jersey: John Wiley & Sons Inc; 2014.