*mathematics*

MDPI

*Article*

# Causality-Driven Efficient Feature Selection for Deep-Learning-Based Surface Roughness Prediction in Milling Machines

**Hyeon-Uk Lee [1,2], Chang-Jae Chun [3] and Jae-Mo Kang [2,*]**

[1] Artificial Intelligence Research Center, Korea Electrotechnology Research Institute (KERI), Changwon-si 51543, Gyeongsangnam-do, Republic of Korea; chopa7@keri.re.kr
[2] Department of Artificial Intelligence, Kyungpook National University, Daegu 41566, Republic of Korea
[3] Department of Artificial Intelligence, Sejong University, Seoul 05006, Republic of Korea; cjchun84@gmail.com
* Correspondence: jmkang@knu.ac.kr

**Abstract:** This paper studies the application of artificial intelligence to milling machines, focusing specifically on identifying the inputs (features) required for predicting surface roughness. Previous studies have extensively reviewed and presented useful features for surface roughness prediction. However, applying research findings to actual operational factories can be challenging due to the additional costs of sensor installations and the diverse environments present in each factory setting. To address these issues, in this paper, we introduced effective features for predicting surface roughness in situations where additional sensors are not installed in the existing environment. These features include feed per tooth, $F_z$; material removal rate, $Q$; and the load information. These features are suitable for use in highly constrained environments where separate sensor installation is not required, making it possible to apply the research findings in various factory environments. Additionally, to efficiently select the optimal subset for surface roughness prediction among subsets formed by available features, we apply causality to the feature selection method, proposing an approach called causality-driven efficient feature selection. The experimental results demonstrate that the features introduced in this paper are quite suitable for predicting surface roughness and that the proposed feature selection approach is more effective and efficient compared to existing selection methods.

**Keywords:** surface roughness; feature selection; deep learning; computer numerical control; milling machine

**MSC:** 68T05



**Citation:** Lee, H.-U.; Chun, C.-J.; Kang, J.-M. Causality-Driven Efficient Feature Selection for Deep-Learning-Based Surface Roughness Prediction in Milling Machines. *Mathematics* **2023**, *11*, 4682. https://doi.org/10.3390/math11224682

Academic Editors: Stefano Carrino, Vicente Rodríguez Montequín, Hatem Ghorbel and Ivana Budinská

Received: 20 October 2023
Revised: 10 November 2023
Accepted: 13 November 2023
Published: 17 November 2023



## 1. Introduction

In the manufacturing industry, many approaches have been taken to improve the competitiveness of factories [1]. Some of them have shown good results, with automation being particularly well utilized in factories. The milling machine, which is widely used in the manufacturing industry and which produces products by cutting workpieces into desired shapes, is a good example of the application of automation by computer numerical control to improve productivity. In addition to improving productivity through the use of automated milling machines, it is also essential to ensure the quality and reliability of the product. In recent times, powered by artificial intelligence, the fourth industrial revolution has had its impact on the manufacturing industry. Artificial intelligence is one of the most powerful methods to improve the competitiveness of factories. Many studies have been conducted to improve productivity more by applying deep learning. In particular, previous studies [2–11] have extensively reviewed and presented features that can be used as inputs to deep learning to predict surface roughness. Surface roughness is a significant index to

evaluate product quality and an indicator of product characteristics that include surface friction and fracture resistance.

Specifically, in recent studies [2,3], surface roughness was predicted using the operating parameters of a milling machine as inputs to deep learning. Study [2] used spindle speed, feed rate, and depth of cut as inputs to the combined model, which was a response surface methodology and genetic algorithm. This work was extended to use down-feed as an additional input with the fully-connected (FC) model in [3]. Although these inputs can be easily and directly obtained from the operating parameters, they are limited in that they do not contain enough information to predict surface roughness. To address these limitations, previous studies expanded and reviewed tools [4] and materials [5,6] as additional inputs along with operating parameters. On the other hand, in works [7–11], additional sensors were installed and used as inputs to acquire more informative data from the milling operations. Works [7,8] installed vision sensors to acquire image data, which were then converted to gray-scale levels to be used as inputs. In works [9–11], different sensors (force, vibration, and acceleration, respectively) were included to acquire more informative information.

Although works [2–11] presented and reviewed many of the available features, their research findings are not easily applicable to the manufacturing industry (factories). Utilizing the features examined in previous research, the problem of additional sensor costs arises, which burdens manufacturing plants striving to improve productivity with this extra expense. Furthermore, the diverse environments in factories pose limitations in replicating the environment where the research findings were obtained into actual factory fields. Even if the challenges related to sensor installation and diverse environments are addressed, the issue of selecting the optimal feature combination for predicting surface roughness remains. The number of possible subsets increases exponentially with the number of all available features, and most of these subsets do not need to be reviewed [12]. Feature selection algorithms have three basic methods: filter, embedded, and wrapper. Typically, the performance of the wrapper selection method is known to be better in the machine learning field [12]. Even with the wrapper selection method, a large search space is still required to determine the optimal subset.

Motivated by all of the aforementioned issues, in this paper we study the available features and feature selection for surface roughness prediction. Specifically, we wish to address two questions: (i) What are the features available without the need for installing any additional sensors? and (ii) How can we reduce the search space when there are many available features? In our design, a critical challenge is in reviewing all the available features when using only the milling machine and comparing the number of searches with the selection function of the wrapper approach. We effectively and efficiently address this challenge by identifying causality among the features and applying it to the selection function. The main contribution of our work is summarized as follows:

- We describe all the available features, in practical situations with no additional sensors installed, that can be used for surface roughness prediction with deep learning.
- Considering causality among the features, we also propose a new and efficient feature selection approach for surface roughness prediction. This scheme reduces the search steps and maintains or improves the accuracy of surface roughness prediction, even when there are many features available.
- Experimental results are presented to demonstrate the superiority and effectiveness of the proposed scheme in terms of selection compared to several baseline selections.

## 2. Available Features for Deep-Learning-Based Surface Roughness Prediction

In this section, we describe two different sets of available features for surface roughness prediction with deep learning: one set involves the conventional features typically used in the previous studies, and the other set involves the new features introduced in this paper for better learning.

### *2.1. Conventional Features: Cutting Parameters Less Relevant to Surface Roughness Prediction*

In practice, a number of factors (such as tool, workpiece, cutting parameters, temperature, coolant, etc.) affect the surface roughness in milling machines in a very complicated way [13]. Thus, a naive and natural approach is to extract the features for surface roughness prediction by considering as many measurable factors as possible. This approach has been taken in the previous studies [2,3]. Specifically, in [2,3], the major operating parameters of computer-aided manufacturing (CAM), usually used to generate work orders (G-code), were adopted as input features for the deep learning algorithms as follows:

- Feed rate $F$ (in mm/min): This feature specifies the tool's moving speed in the cutting direction.
- Spindle speed $S$ (in r/min): This feature specifies the rotation speed of the spindle motor.
- Depth $D$ or width $W$ of cut (in mm): This feature specifies depth or width of the material cut per pass of the tool.

Although these conventional features (i.e., the set of $F$, $S$, $D$, and $W$) can be directly obtained from the CAM, use of only these features has several major limitations on surface roughness prediction with deep learning. First of all, the ability to predict the surface roughness with the conventional features is very poor (regardless of using the state-of-the-art techniques such as deep learning) because such features were heuristically and empirically chosen. Even in theory, these features are less relevant and only indirectly affect to the surface roughness prediction. To address this critical issue, therefore, the conventional features should be used together with other features that can lead to better learning or prediction performance (this will be explained in the next section in detail).

The second critical issue is that, the feature extraction approach of the previous studies requires installation of several devices (such as sensors) on the milling machines, which is costly and also inconvenient in practice. Additionally, the variances of the extracted features would be very high due to the inherent randomness in the state of the machines, and this would unavoidably lead to unstable prediction results.

### *2.2. New Features for Better Learning: Load Information and Cutting Parameters More Relevant to Surface Roughness Prediction*

In order to address all the critical issues resulting from using only the conventional features, in this paper, we introduce new effective features that result in much better prediction performance with deep learning, particularly when combined with the conventional features. Detailed explanations are given as follows.

#### 2.2.1. Cutting Parameters More Relevant to Surface Roughness Prediction

According to the metal cutting theory, there are features that are more direct and relevant to the surface roughness prediction. Inspired by this, in this paper, such features are additionally exploited for surface roughness prediction with deep learning, which are as follows [13]:

- Feed per tooth $F_z$ (in mm/tooth): This feature specifies the cutting speed on the tooth of the tool and is given by $F_z = FS^{-1}z^{-1}$, where $z$ is the number of teeth of the tool.
- Material removal rate $Q$ (in $\text{mm}^3$/min): This feature specifies the rate of material removal and is given by $Q = FDW$.

The proposed cutting parameter (more relevant) differs from conventional cutting parameters (less relevant) in that it encompasses features related to the motion of rotation, direction of movement, and material removal of the cutting tool. Surface roughness is an indicator that evaluates the structures generated during the interaction between the cutting tool and the workpiece in the cutting process. Thus, these parameters are more intuitively associated with the surface roughness resulting from the relative motion of the tool and workpiece; therefore, they assist in better learning for predicting surface roughness.

Interestingly, it turns out that the features being more relevant to the surface roughness prediction are actually given by combinations of the conventional features.

#### 2.2.2. Load Information

From the milling machine, we can also extract further useful information, called the load information, about the load resistance that occurs during cutting. The load information is given in the form of time series, for which statistics (i.e., a more abstract and compact version of the load information) are typically used for surface roughness prediction. Inspired by this, in this paper, we also propose to use such statistics as part of the input features for the deep learning algorithms. Specifically, let us denote the load information by $\mathbf{x}(n) \in \mathcal{R}^L$ for $n = 1, \cdots, N$. We can extract the following eight useful statistics about the load information as features:

(1) The load information itself: $\mathbf{x}(n)$
(2) Average value: $V_{avg} = \frac{1}{N}\sum_{n=1}^{N}\mathbf{x}(n)$
(3) Maximum value: $V_{max} = \max_{n=1,\cdots,N}\{\mathbf{x}(n)\}$
(4) Minimum value: $V_{min} = \min_{n=1,\cdots,N}\{\mathbf{x}(n)\}$
(5) Variance: $V_{var} = \frac{N\sum_{n=1}^{N}\mathbf{x}^2(n)-\left(\sum_{n=1}^{N}\mathbf{x}(n)\right)^2}{N(N-1)}$
(6) Skewness: $V_{skew} = \frac{N}{(N-1)(N-2)}\sum_{n=1}^{N}\left(\frac{\mathbf{x}(n)-V_{avg}}{\sqrt{V_{var}}}\right)^3$
(7) Kurtosis: $V_{kur} = \frac{\sum_{n=1}^{N}(\mathbf{x}(n)-V_{mean})^4}{n\times {V_{var}}^2}$
(8) Coefficient variation: $V_{cv} = \frac{\sqrt{V_{var}}}{V_{avg}}$

One big practical advantage of using the load information is that there is no need to additionally install sensors on the machines. This is in sharp contrast to the case of learning solely with the conventional features. In order to achieve the best prediction performance and draw useful engineering insights into which features critically affect the prediction performance, the remaining question is how to select the best combination across the conventional and newly introduced features. This important question is answered in the next section.

## 3. Causality-Driven Efficient Feature Selection for Deep-Learning-Based Surface Roughness Prediction in Milling Machines

The exhaustive selection approach of the wrapper selection method is adopted naturally and simply to select the best feature set. This search guarantees the selection of the best feature set. However, it is a very difficult task because it requires a large search space due to all possible subsets of the features. All possible subsets can be easily calculated as $2^m$, where $m$ represents all available features. In our case, since empty sets are not allowed in deep learning, there are 16,383 possible subsets calculated using $2^m - 1$. However, there are many subsets that do not need to be evaluated as they are relatively less important. Thus, it is essential to reduce the number of searches [12].

Typically, in order to address this issue, many studies have heuristically organized the subsets [2–11]. However, these approaches lead to other problems, such as bias due to unfounded standards. Thus, to address these issues and effectively reduce the number of searches, we present the causality-driven efficient feature Selection for deep learning-based surface roughness prediction in milling machines in Algorithm 1.

**Algorithm 1:** Proposed causality-driven efficient feature selection for deep-learning-based surface roughness prediction in milling machines.

**Input:** All of available subsets pre-categorized $\mathcal{C}$, $\mathcal{V}$, and $\mathcal{X}$
**Output:** Sub-optimal set of features $X^*$, and Predict accuracy $\varphi^*$
**begin**
  **Function** `Selection(`$X_1$, $X_2$`):`
    initialization $X_3 = \{\phi\}$, $\varphi_3 = 0$
    **while** *True* **do**
      `/* Forward step */`
      **if** $X_3 = \{\phi\}$ **then**
        $X_{1'} = X_1$
      **else**
        $X_{1'} = X_1 - X_3$
      **end**
      **for** $i = 1$ *to* $|X_{1'}|$ **do**
        $X_{a-i} = (x_i \in X_{1'}) \cup X_2 \cup X_3$
      **end**
      $X_A = [X_{a-1}, \cdots, X_{a-i}]$
      $X_{A*} = \underset{x \in X_A}{argmax}(Evaluation(x, M))$
      $\varphi_{A*} = Evaluation(X_{A*}, M)$
      $X_{a*} = X_{A*} - X_2$
      **if** $\varphi_{A*} > \varphi_3$ **then**
        $\varphi_3 = \varphi_{A*}$
        $X_3 = X_{a*}$
      **else**
        **if** $\varphi_{A*} < \varphi_3$ **then**
          $\varphi_3 = \varphi_3$
          $X_3 = X_3$
        **else**
          **if** $\varphi_{A*} = \varphi_3$ **then**
            **return** $X_2 \cup X_3$, $\varphi_3$
          **end**
        **end**
      **end**
      `/* Backward step */`
      **if** $1 \neq |X_{a*}|$ **then**
        **for** $j = 1$ *to* $(|X_{a*}|)$ **do**
          $X_{b-j} = X_{a*} - (x_j \in X_{a*})$
          $X_{b-j} = X_{b-j} \cup X_2 \cup X_3$
        **end**
        $X_B = [X_{b-1}, \cdots, X_{b-j}]$
        $X_{B*} = \underset{x \in X_B}{argmax}(Evaluation(x, M))$
        $\varphi_{B*} = Evaluation(X_{B*}, M)$
        $X_{b*} = X_{B*} - X_2$
        **if** $\varphi_{B*} > \varphi_3$ **then**
          $\varphi_3 = \varphi_{B*}$
          $X_3 = X_{b*}$
        **else**
          **if** $\varphi_{B*} < \varphi_3$ **then**
            $\varphi_3 = \varphi_3$
            $X_3 = X_3$
          **else**
            **if** $\varphi_{B*} = \varphi_3$ **then**
              **return** $X_2 \cup X_3$, $\varphi_3$
            **end**
          **end**
        **end**
      **end**
    **end**
  **end**
  $\mathcal{C}' = Selection(\mathcal{C}, \phi)$
  $\mathcal{V}' = Selection(\mathcal{V}, \mathcal{C}')$
  $\mathcal{X}' = \mathcal{C}' \cup \mathcal{X}$
  $X_{sub-opt} = \{\mathcal{C}', \mathcal{V}', \mathcal{X}'\}$
  $X^* = \underset{x \in X_{sub-set}}{argmax}\ (Evaluation(x, M))$
  $\varphi^* = Evaluation(X^*, M)$
**end**

### *3.1. Proposed Algorithm*

Before drawing the algorithm, we organize all available features into $\mathcal{C}$, $\mathcal{V}$, and $\mathcal{X}$ according to the data source and type. A summary of the available features mentioned earlier is as follows:

- $\mathcal{C} = \{F, S, D, W, F_z, Q\}$
- $\mathcal{V} = \{V_{avg}, V_{max}, V_{min}, V_{var}, V_{skew}, V_{kur}, V_{cv}\}$
- $\mathcal{X} = \{\mathbf{x}(n) : n = 1, \cdots, N\}$

The organized features can be divided into cause ($\mathcal{C}$) and effect ($\mathcal{V}, \mathcal{X}$) based on the conditions of data acquisition, which allows for the definition of causality. The causality established from this division is used in the proposed algorithm to effectively reduce the search steps and eliminate relatively irrelevant, redundant, and noisy subsets in the prediction of surface roughness. The proposed scheme for applying causality consists of two parts: (i) selection function and (ii) selection strategy. A detailed explanation of each part follows.

#### 3.1.1. The Selection Function in the Proposed Algorithm

In the selection function, we employed a step-wise selection approach with two steps (forward and backward) in the wrapper selection method. This approach requires more searches than other approaches with one step; however, previous studies have shown that a near-optimal subset can be selected using this approach [12]. In order to take together the organized features with different causalities, the selection function in the proposed algorithm takes two types of inputs: the set of variables $X_1$ and the set of constants $X_2$, where $X_1$ consists of the features that need to be selected and $X_2$ consists of the fixed or previously selected features.

In detail, during the forward step, each feature of $X_1$ is added to $X_2$ to generate the forward candidate subset $X_A$. The components of $X_A$ are then used separately as inputs to the deep learning model to determine the valid subset $X_{a*}$, which has the most effective features at each forward step. Afterward, in the backward step, all possible subsets that can be created by removing one element from $X_{a*}$ are added to $X_2$ to generate the backward candidate subset $X_B$. The components of $X_B$ are also used separately as inputs to the deep learning model, just as in the forward step, and valid subset $X_{b*}$ is determined at each backward step. The two steps are repeated, retaining the effective feature subset $X_3$ at each step, until there is no improvement in the accuracy $\varphi_3$ of surface roughness prediction.

The selection function presented faces a problem when generating $X_B$ during the first repeat, and that needs to be resolved. As there is only one valid element determined from the forward step on the first repeat, the backward step cannot create an appropriate $X_B$. To address this issue, the selection function does not implement the backward step during the first repeat but only implements it when there are two or more valid elements selected from the forward step.

#### 3.1.2. The Selection Strategy of the Proposed Algorithm

In the selection strategy, we select a combination of features by taking the organized feature sets (i.e., $\mathcal{C}$, $\mathcal{V}$, and $\mathcal{X}$) sequentially as inputs with the aim of applying causality. The selection function initially takes $\mathcal{C}$ and the empty set as inputs and determines $\mathcal{C}'$, which is a valid feature in $\mathcal{C}$. Here, $\mathcal{C}$ is considered as the set of variables $X_1$, and the empty set is considered as the set of constants $X_2$. Then, $\mathcal{V}'$ and $\mathcal{X}'$ are determined by sequentially taking the remaining organized features ($\mathcal{V}$ and $\mathcal{X}$) as $X_1$. At this time, the previously selected $\mathcal{C}'$ is taken into $X_2$ to apply causality. Finally, to select the sub-optimal set of features $X^*$ from the derived features ($\mathcal{C}'$, $\mathcal{V}'$, and $\mathcal{X}'$), we use a function with arguments of the maxima as follows:

$$X^* = \underset{x \in X_{sub-opt}}{argmax} \left( Evaluation(x, M) \right) \tag{1}$$

where $M$ is the deep learning model for surface roughness prediction and $X_{sub-opt}$ contains the candidate set of sub-optimal $\mathcal{C}'$, $\mathcal{V}'$, and $\mathcal{X}'$. This selection strategy has the advantage of being able to prioritize and control the organized features.

## 4. Experimental Verification

In this section, we describe the materials and experimental results used to verify the newly introduced features and the proposed algorithm.

### *4.1. Materials*

#### 4.1.1. Benchmark Datasets

To validate the effectiveness of the proposed features in improving the accuracy of surface roughness prediction, we employed two different benchmark datasets. (i) Alam et al. (2023) [2] obtained the material by using a high-speed milling machine to process titanium (Ti-6Al-4V) with four flutes and 2 mm uncoated carbide tools to conduct experiments for predicting and optimizing surface roughness. The machining was conducted with 20 experimental runs in a straight-line configuration, varying the values of $S$, $F$, and $D$. Surface roughness measurements were taken using the Surftest SJ-201 (Mitutoyo, Kanagawa, Japan), and an average of three readings was obtained. (ii) Dubey et al. (2022) [6] obtained experimental data using a CNC lathe to machine Al7075 with Widia's CNMG120408 (Widia, Koenigsee-Rottenbach, Germany) Tungsten Carbide insert tools for surface roughness prediction using machine learning approaches. They conducted 27 experimental runs, varying the conditions of $S$, $F$, and $D$ into three categories each for machining. Surface roughness was measured using the Surftest SJ-201 (Mitutoyo, Kanagawa, Japan) and the results were provided.

#### 4.1.2. Experimental Data

A medium-sized milling machine with aluminum and end-mill tools, a configuration widely used in the manufacturing industry, is selected for the trials. Specifically, the milling machine model is Mynx 5400 (DN Solutions, Changwon-si, Republic of Korea) from DN Solutions. Al6061 is chosen as the workpiece material, and Alu-cut 8 mm from YG-1 is chosen as the end mill. The tool path for the machining operation is a straight line and is implemented through 260 different work orders consisting of combinations of components in $\mathcal{C}$. During the operation, the load information is acquired using the FANUC Open CNC API Specifications library (FOCAS) and Open Platform Communications Unified Architecture (OPC-UA) [14]. After each operation, the surface roughness is measured using the MITUTOYO SJ-210 surface roughness tester, and the average value is calculated based on four to five measurements taken at different locations on the workpiece. This is performed to prevent observation errors and ensure the accuracy of the measurement.

#### 4.1.3. Model

The deep learning model for predicting surface roughness has been developed using the environment of Python (version 3.8) and Tensorflow (version 2.5). To select the feature set with the highest accuracy of surface roughness prediction in the proposed algorithm, two models were used depending on the type of data: one was a fully connected (FC) model with five hidden layers, each containing 150 nodes; the other was a combined model, consisting of an FC model with five hidden layers and a recurrent neural network (RNN) with four hidden layers. In the combined model, the RNN was concatenated after the fourth hidden layer of the FC model. For general experimental results, the hyperparameters of deep learning were determined using techniques widely used for similar problems: (i) a rectified linear unit (ReLu) at each layer and a linear unit at the last layer; (ii) optimizer: root mean square propagation (RMSprop); (iii) $k$-fold where $k$ was 10; (iv) normalization [15]; and (v) loss function: mean squared error (MSE).

### *4.2. Result and Discussion*

In order to better evaluate and understand the effectiveness and performance of the proposed scheme, we compared it with the following four baselines of wrapper selection [12]:

- Baseline 1: A sequential selection approach that used only forward steps starting with an empty state.
- Baseline 2: A recursive selection approach that used only backward steps, starting with a full consideration of the features.
- Baseline 3: A step-wise selection approach that was a combination of the above two selected approaches and was the basis of our algorithm.
- Exhaustive Search: An exhaustive selection approach that used all the possible subsets of the features.

To ensure a fair and unbiased comparison of our proposed method across various datasets, we apply each dataset to two evaluations: (i) The effect of the cutting parameters (more relevant) for benchmark data and experimental data. (ii) The efficiency of the proposed algorithm and the effect of load information for experimental data. Since the benchmark data cannot provide load information, it is solely used for evaluating the effect of cutting parameters.

#### 4.2.1. The Effect of the Proposed More Relevant Cutting Parameters

Figures 1 and 2 show the results of conventional feature selections (baselines) for benchmark data, and Figure 3 shows the results for experimental data. For Figure 1, all conventional feature selections select the optimal subset ($Q$, $F_z$, $S$, $D$) for the highest surface roughness prediction, achieving a prediction accuracy of 96.36%. In particular, the results show that Baselines 1 and 3 prioritize the selection of $Q$ and $F_z$ in the forward steps to improve the accuracy of surface roughness prediction. In the case of Figure 2, most conventional feature selection methods achieve optimal results, whereas Baseline 1 does not. The optimal subset of this case is $F_z$, $D$, and $F$, with a prediction accuracy of 91.62%. In this case, which uses a different dataset from the case in Figure 1, $Q$ and $F_z$ are also selected as the most important features in the forward steps of Baselines 1 and 3 to improve the prediction accuracy of surface roughness. This finding is consistent with the results of the case in Figure 1. In the case in Figure 3, most conventional feature selection methods still achieve optimal results, but recursive feature selection does not. The optimal subset of this case is $F_z$, $W$, $D$ and $Q$, with a prediction accuracy of 84.91%. In this case, $F_z$ is selected as the most important feature in the forward steps of Baselines 1 and 3 to improve the prediction accuracy of surface roughness.

In all of datasets prepared, our proposed cutting parameters ($Q$ and $F_z$) are prioritized over other features to improve the prediction accuracy of surface roughness. This discovery validates the effectiveness of our proposed features in improving the accuracy of predicting surface roughness.

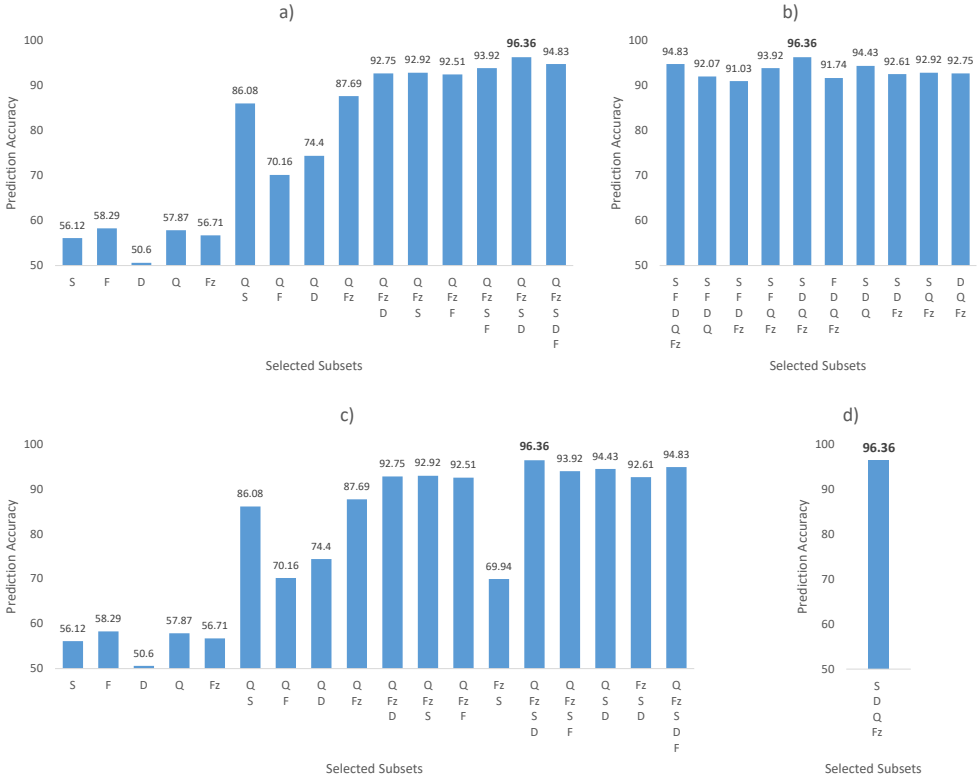


**Figure 1.** The Results of feature selection using benchmark data [2], where bold numbers indicate the highest prediction accuracy with the best subset: (**a**) Baseline 1 (Sequential), (**b**) Baseline 2 (Recursive), (**c**) Baseline 3 (Step-wise), and (**d**) Exhaustive.

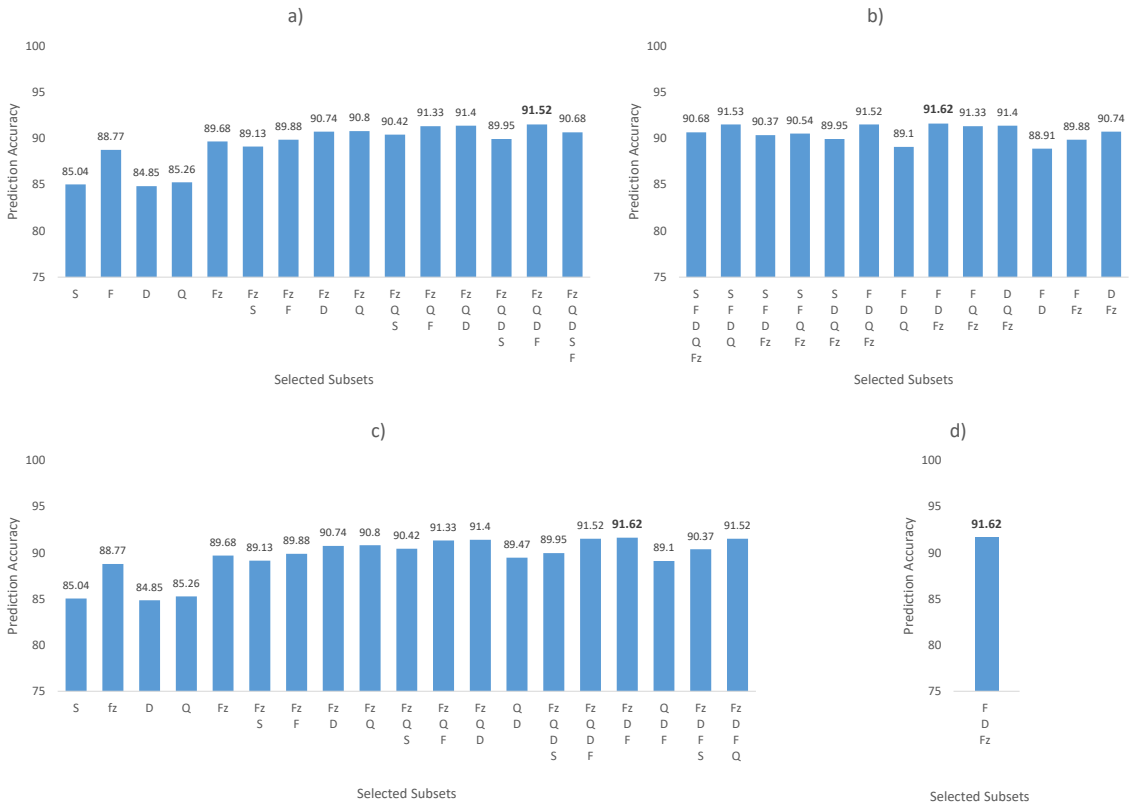


**Figure 2.** The results of feature selection using benchmark data [6], where bold numbers indicate the highest prediction accuracy with the best subset : (**a**) Baseline 1 (Sequential), (**b**) Baseline 2 (Recursive), (**c**) Baseline 3 (Step-wise), and (**d**) Exhaustive.

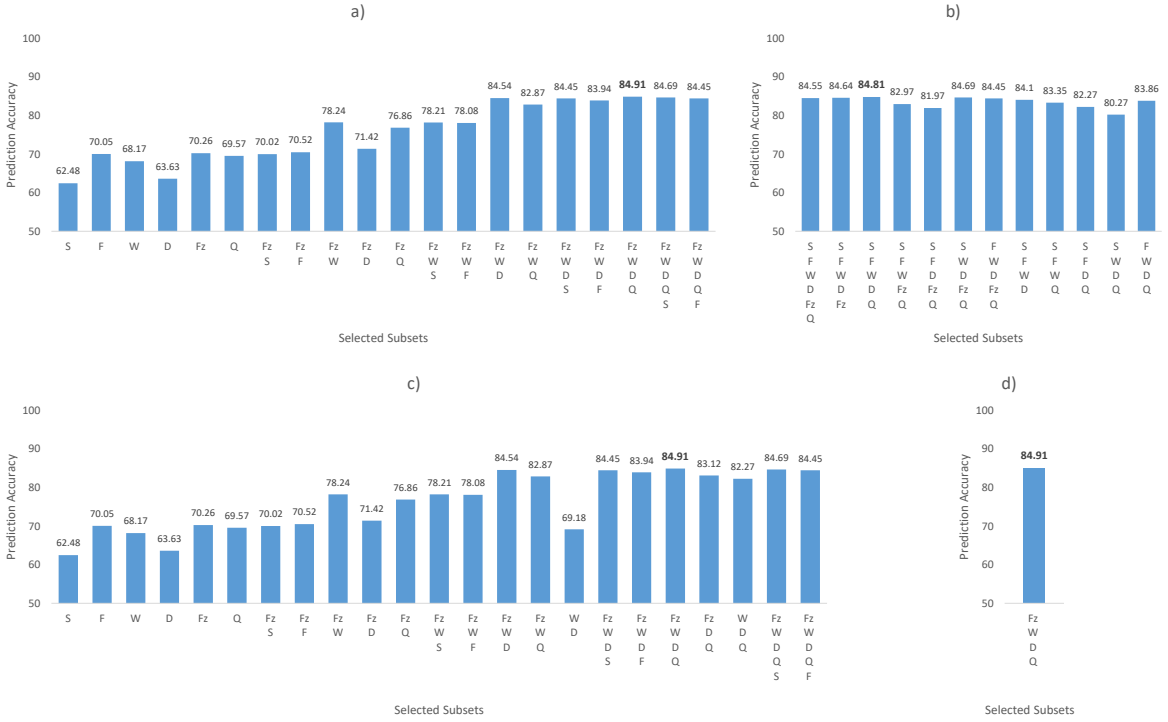


**Figure 3.** The results of feature selection using experimental data, where bold numbers indicate the highest prediction accuracy with the best subset: (**a**) Baseline 1 (Sequential), (**b**) Baseline 2 (Recursive), (**c**) Baseline 3 (Step-wise), and (**d**) Exhaustive.

#### 4.2.2. The Efficiency of Proposed Scheme and the Effect of Load Information

To verify the efficiency of our selection algorithm, we organized the experimental data in a causality and applied them to the proposed algorithm. Since $\mathcal{V}$ and $\mathcal{X}$ are different forms of load information, we evaluated the efficiency of our algorithm in two scenarios: scenario 1 (using $\mathcal{C}$ and $\mathcal{V}$) and scenario 2 (using $\mathcal{C}$, $\mathcal{V}$, and $\mathcal{X}$, considering all of them).

Figure 4 illustrates the efficiency of our proposed causality-driven efficient feature selection algorithm. Baseline 2 is generally the preferred selection method because it is the fastest. However, as demonstrated by our experimental results (in Figures 3 and 4), Baseline 2 does not guarantee the optimality of the selected subset. In comparison, our proposed selection algorithm can achieve at least as good or even better feature selection results than Baseline 3, even with fewer search steps. Furthermore, our proposed selection can achieve better results with fewer search steps compared to Baseline 1, which exclusively uses forward steps. The improvement effect of load information on surface roughness prediction accuracy can be verified by using Figures 3 and 4. Utilizing statistical features of load information increases the surface roughness prediction accuracy by 2.54% (from 84.91% to 87.45%), whereas employing raw data improves the accuracy by 6.93% (from 84.91% to 91.84%). Interestingly, accuracy is higher on most occasions when using the load information in its raw form compared to its statistical form. The highest accuracy of 91.84% was achieved when using the subsets $D$, $W$, $F_z$, $Q$, and $\mathcal{X}$ as inputs to predict surface roughness. The highest accuracy of 91.84% is achieved when utilizing a subset (features: $D$, $W$, $F_z$, $Q$, and $\mathcal{X}$) for predicting surface roughness. Our experimental results demonstrate that load information can be used as an effective feature to improve the accuracy of surface roughness prediction in situations where sensor installation is limited.

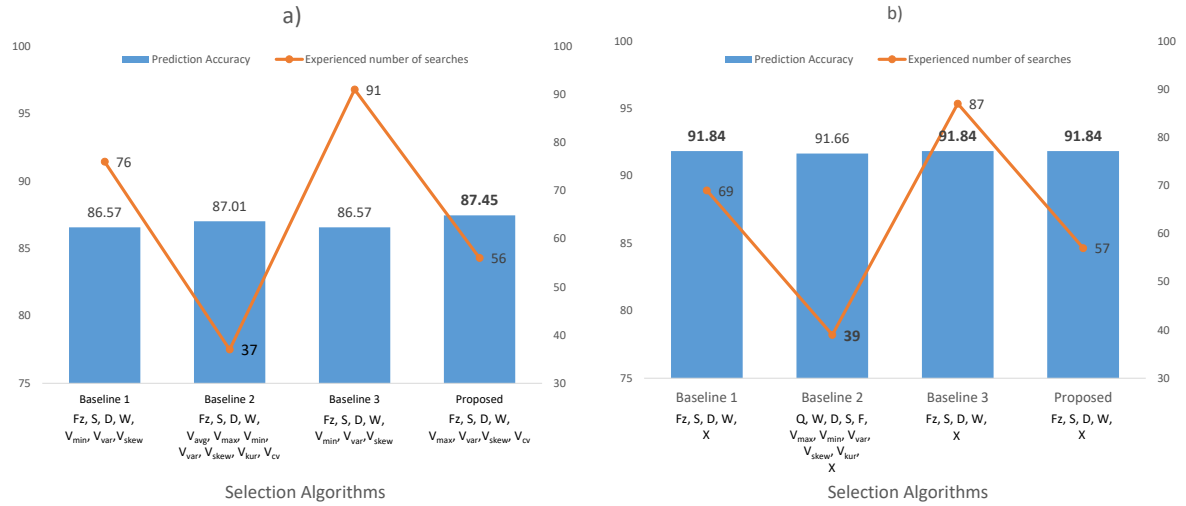


**Figure 4.** The feature selection results of experimental data with causality applied, where bold numbers indicate the highest prediction accuracy with the best subset: (**a**) Utilized organized features $\mathcal{C}$, $\mathcal{V}$, where all possible subsets are 8191. (**b**) Utilized organized features $\mathcal{C}$, $\mathcal{V}$, and $\mathcal{X}$, where all possible subsets are 16,384.

## 5. Conclusions and Future Work

In this paper, we introduced new, effective features that resulted in much better prediction of surface roughness using deep learning and proposed causality-driven efficient feature selection for deep-learning-based surface roughness prediction in milling machines. Experimental results using both benchmark data and our data showed that our proposed method was more effective and efficient than the conventional method.

Our research aims to develop an algorithm that accurately predicts surface roughness under various machining conditions and selects the most important features that contribute to the surface roughness prediction. One limitation of our research is that the load information can be different even if the machining conditions do not change (e.g., machine aging, tool and material mismatch, etc.). In this situation, cutting parameters cannot be used to predict surface roughness, so it is difficult to directly apply our research results. In future works, two main directions are suggested. One is to develop a method to predict surface roughness using only load information, considering the situations mentioned earlier. The other is to propose machining conditions that maintain the surface roughness standard of the workpiece using load information, as well as predicting surface roughness using load information.

**Author Contributions:** Conceptualization, H.-U.L. and J.-M.K.; methodology, H.-U.L.; software, H.-U.L.; validation, C.-J.C. and J.-M.K.; formal analysis, C.-J.C. and J.-M.K.; investigation, H.-U.L.; resources, H.-U.L.; data curation, H.-U.L.; writing—original draft preparation, H.-U.L.; writing—review and editing, C.-J.C. and J.-M.K.; visualization, H.-U.L.; supervision, J.-M.K.; project administration, J.-M.K.; funding acquisition, H.-U.L. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was supported by Korea Electrotechnology Research Institute (KERI) primary research program through the National Research Council of Science & Technology (NST) funded by the Ministry of Science and ICT (MSIT) (No. 20A01014, 23A01015).

**Data Availability Statement:** The benchmark data presented in this study (used in Figure 1) is openly available in Table 4 at https://doi.org/10.1016/j.rineng.2022.100873, with the reference number [2]. The benchmark data presented in this study (used in Figure 2) is openly available in Table 3 at https://doi.org/10.3390/lubricants10050081, with the reference number [6]. The data presented in this study (used in Figures 3 and 4) is available on request from the corresponding author. This data is not publicly available due to being funded by the Ministry of Science and ICT (MSIT) and requiring institutional approval.

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

## References


1. Zhong, R.Y.; Xu, X.; Klotz, E.; Newman, S.T. Intelligent manufacturing in the context of industry 4.0: A review. *Engineering* **2017**, *3*, 616–630. [CrossRef]
2. Alam, S.T.; Tomal, A.A.; Nayeem, M.K. High-speed machining of Ti–6Al–4V: RSM-GA based optimization of surface roughness and MRR. *Res. Eng.* **2023**, *17*, 100873. [CrossRef]
3. Danh, B.T.; Tung, L.A.; Danh, T.H.; Tan, T.M.; Tuan, T.K.; Van Phuong, D.; Ngoc, N.D. Studying Optimal Set of Input Parameters for CBN Grinding Aluminum 6061T6 on CNC Milling Machine. In Proceedings of the International Conference on Engineering Research and Applications, Thai Nguyen, Vietnam, 1–2 December 2022; Springer: Berlin/Heidelberg, Germany, 2023; pp. 895–902.
4. Sizemore, N.E.; Nogueira, M.L.; Greis, N.P.; Davies, M.A. Application of machine learning to the prediction of surface roughness in diamond machining. *Procedia Manuf.* **2020**, *48*, 1029–1040. [CrossRef]
5. Kesarwani, S.; Kharwar, P.K.; Verma, R.K. Artificial Neural Network (ANN) for Parametric Appraisal and Milling Efficiency Evaluation of Carbon/Epoxy Nanocomposites. In *Advances in Simulation, Product Design and Development*; Springer: Berlin/Heidelberg, Germany, 2023; pp. 167–179.
6. Dubey, V.; Sharma, A.K.; Pimenov, D.Y. Prediction of surface roughness using machine learning approach in MQL turning of AISI 304 steel by varying nanoparticle size in the cutting fluid. *Lubricants* **2022**, *10*, 81. [CrossRef]
7. Babu, A.R. A Machine Learning Approach for Prediction of Surface Roughness from the Images of Machined Components. In *Advances in Forming, Machining and Automation*; Springer: Berlin/Heidelberg, Germany, 2023; pp. 405–416.
8. Karthikeyan, S.; Subbarayan, M.; Beemaraj, R.K.; Sivakandhan, C. Computer vision-based surface roughness measurement using artificial neural network. *Mater. Today Proc.* **2022**, *60*, 1325–1328.
9. Balasubramanian, A.N.; Yadav, N.; Tiwari, A. Analysis of cutting forces in helical ball end milling process using machine learning. *Mater. Today Proc.* **2021**, *46*, 9275–9280. [CrossRef]
10. Lin, W.J.; Lo, S.H.; Young, H.T.; Hung, C.L. Evaluation of deep learning neural networks for surface roughness prediction using vibration signal analysis. *Appl. Sci.* **2019**, *9*, 1462. [CrossRef]
11. Möhring, H.C.; Eschelbacher, S.; Georgi, P. Machine learning approaches for real-time monitoring and evaluation of surface roughness using a sensory milling tool. *Procedia CIRP* **2021**, *102*, 264–269. [CrossRef]
12. Kumar, V.; Minz, S. Feature selection: A literature review. *SmartCR* **2014**, *4*, 211–229. [CrossRef]
13. Stephenson, D.A.; Agapiou, J.S. *Metal Cutting Theory and Practice*; CRC Press: Boca Raton, FL, USA, 2018.
14. Martins, A.; Lucas, J.; Costelha, H.; Neves, C. Developing an OPC UA Server for CNC Machines. *Procedia Comput. Sci.* **2021**, *180*, 561–570. [CrossRef]
15. Zain, A.M.; Haron, H.; Sharif, S. Prediction of surface roughness in the end milling machining using Artificial Neural Network. *Expert Syst. Appl.* **2010**, *37*, 1755–1768. [CrossRef]