MEASUREMENT AND CONTROL

Intelligent monitoring of linear stages with ensembles of improved LeNET DCNN and random forest classifiers

Mustafa Demetgul[3] ![ORCID], Yicheng Zhao[1],
Ibrahim Nur Tansel[2] and Jurgen Fleischer[1]

Abstract

The linear stages are the most critical component of machine tools and additive manufacturing equipment. The accuracy of the linear stages directly affects the quality of the parts produced. Misalignment is a common problem in the linear stages. This paper presents a sensorless approach for detecting misalignment by monitoring the motor current. A linear stage was designed to simulate various angular misalignment problems between the ball screw and the motor shaft. The sensorless current-based method monitored the motor current at the Programmable Logic Controller (PLC) to detect the misalignment of the linear stage. Different forces were applied to the linear stage under different misalignment conditions. The acquired signal was processed using Continuous Wavelet Transform (CWT). The Lenet DCNN (Deep Convolutional Neural Network) model structure was improved by hyper-parametertuning and ensemble. The ensemble method combined the Convolutional Neural Network (CNN) model with a random forest (RF) classifier. The developed anomalydetection system was trained when different forces, with and without misalignment, were applied. The results showed that the proposed method was feasible for estimating the misalignment even when different external forces were applied to the linear stage.

Keywords

Sensorless monitoring, improved LeNet5, hyperparameter tuning, ensemble, misalignment

Date received: 28 June 2024; accepted: 16 January 2025

Introduction

Machine tools, additive manufacturing equipment and many other machines use linear stages. The accuracy of the linear stages determines the quality of the parts produced. Early detection and correction of misalignment problems reduces quality control problems and machine downtime.[1,2]

Linear stages are driven by either a linear motor or a rotary motor aligned with a ball screw. Misalignment of the motor with the ball screw is a common problem.[3] It can be easily measured using a straightedge or dial indicator method when the machine is not running. Misalignment is observed when the motor and ball screw axes are not on the same center line. The edge load caused by a misaligned shaft will damage the bearing. In industry, the axis misalignment of machines is measured with a laser axis misalignment device or mechanically. The laser is expensive, and the mechanical one has low sensitivity. In addition, for these operations, the machine must be stopped, and measurements must be made at certain intervals. This is an impractical and expensive method. To avoid this, the alignment condition must be monitored. Since there are many machines in a factory, this technique is both cheap and offers the opportunity to intervene immediately in the problems of the machines. Machine learning algorithms also allow the machine to report its own problems.

In recent years, various sensors and signal processing methods have been introduced to detect misalignment. Misalignment can be detected by condition monitoring

[1]Institute of Production Science (wbk), Karlsruhe Institute of Technology, Rinheimer Querallee, Karlsruhe, Baden-Württemberg, Germany
[2]Mechanical and Materials Engineering Department, Florida International University, Miami, FL, USA
[3]Institute of Applied Informatics and Formal Description Methods (AIFB), Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany

Data Availability Statement included at the end of the article

Corresponding author:
Mustafa Demetgul, Institute of Applied Informatics and Formal Description Methods (AIFB), Karlsruhe Institute of Technology, AIFB P.O. Kollegiengebäude am Kronenplatz (Bldg. 05.20), Kaiserstr. 89, 76133 Karlsruhe, Germany.
Email: mdemetgul179@gmail.com

systems using vibration,[4] torque,[5] thermal imaging[6] or acoustic sensors. However, they are not widely used in industry because of their disadvantages. The sensors themselves are expensive and need to be installed on the linear table. The expertise of the installer determines the performance of these monitoring systems.[7] In addition to these disadvantages, the installation of additional sensors in the production environment is not desirable. Sensorless misalignment monitoring systems don't have any of the above disadvantages. The sensorless misalignment monitoring system proposed in this paper monitors the current signal of the control loop of the feed drive via PLC software. Compared to sensorbased methods, the cost of the proposed misalignment monitoring system is significantly lower.[8]

Signal processing methods are required to extract the information from the time domain data. The time domain information is lost when the frequency domain methods are used. Considering this limitation, time-frequency analysis methods such as continuous wavelet transform or short-time Fourier transform are preferred[9] to accommodate the signals at different feed rates.

Traditional machine learning algorithms have been used to diagnose manufacturing processes. The features need to be artificially extracted for traditional machine learning algorithms, such as Support Vector Machine (SVM)[10] and K-nearest neighbors algorithm (kNN).[11] Although the traditional algorithms are able to detect the machine anomalies, the feature extraction requires experts to select the most important features in the data. The Convolutional Neural Networks are able to handle the feature extraction by using multiple convolution and pooling layers.

Some simple structures of CNN such as LeNet can extract the feature and classify handwritten numbers with only five layers.[12] Therefore, an expert who understands the feature of different machine errors is no longer needed.[13] The performance of CNN can be further improved by using various techniques, such as hyperparameter tuning[14,15] and ensemble method. Ensemble with simple traditional machine learning algorithms such as random forest is a much more efficient method than combining multiple neural networks.[16,17] RF models have become popular due to their exceptional accuracy, capability to manage complex datasets, and strong resistance to overfitting. These techniques increase the accuracy of the CNN and the reliability of the intelligent fault diagnosis system.[18,19]

The performance of machine learning algorithms depends on the training data provided. Well-balanced data should be provided in each category. In industry, it's not easy to collect data in each category because machine failures do not occur often. Therefore, the data of machine anomaliesare much less than the data collected under healthy operating conditions. Imposing different machine anomaliesto obtain data may damage the machine.[20] To overcome the challenge of unbalanced

data in industry, transfer learning has been used for fault diagnosis.[21–23] Transfer learning is the use of knowledge gained from one application in another application.[24,25]

In this study, experiments were conducted with a linear stage. Ball screws are widely used because of their high efficiency and long service life. Experimental data were collected when there were different angular misalignments between the ball screw and the motor. Different forces were applied to the linear stage at each misalignment condition. The current signal was acquired by the PLC software and then processed with continuous wavelet transform. A CNN model (LeNet) was trained to classify the healthy machine state and different levels of misalignment on the vertical and horizontal axis. Then, hyperparameter tuning and ensemble method were used to improve the accuracy of the CNN model. Based on the CNN model for misalignment detection, transfer learning theory was applied to train a new CNN model for detecting different forces applied to the table. Then, the misalignment is combined with different forces on the table. Transfer learning was applied to this more challenging error detection task.

Related studies

Many studies have been carried out in the field of machine monitoring and feed axis in recent years. Li et al.[26] propose a method for predicting remaining useful life using partial domain adaptation to handle incomplete target data, improving prediction accuracy despite missing information. Chen et al.[27] explored neuromorphic computing for contactless machine fault diagnosis, developing a method for rapid detection across different machines.

One commonly used sensor for misalignment fault detection is the vibration sensor because a misaligned shaft will cause stronger vibration than an aligned one. Guan et al. established dynamic models for angular misalignment and offset misalignment. They detected them by using the spectrum of the vibration signal.[28] Kumar and Kumar used the vibration signal to find the unbalance and the misalignment of the shaft at different speeds and differentiate them with the help of linear discriminant analysis.[29] Wang et al. analyzed the vibration signal of a dual-rotor system with time-frequency techniques. They showed that the time-frequency method can detect misalignment.[30] The acoustic emission (AE), which is designed for static structure has also been employed for misalignment detection. Caso et al. studied misalignment monitoring by using AE sensors. They showed that AE sensors have a significant advantage at low speeds and don't need the complex signal processing needed for the accelerometers.[7] Since sensors are expensive and difficult to assemble, soft sensor techniques are also used. But sometimes it does not behave like a real system.

And there are problems such as quality of data, variable and feature selection, model selection and model construction, soft sensor maintenance.[31]

The stator current signal[30] and rotor current signal[10] have also been used for fault diagnosis. This method has been used to detect the gearbox fault of wind turbines based on current signals collected from the generator. Corne et al. studied the stator current signature of the misalignment in the frequency domain on an 11 kW induction machine.[32]

Most commonly used sensorless approaches use motor current. The current signal is needed for the control loop of the servo motor. Thus it's available for every servo motor without an extra sensor. Plapper and Weck designed a system to monitor the backslash and pitting of the linear stages by using the internal control signals.[33]

Not only current signal but other built-in sensors have also been employed to monitor the machine's condition. Verl et al.[34] quantified and localized the wear of the feed drive using the SCAM algorithm which used positioning and velocity signals. Putz et al.[35] combined positioning and current signals. They used Choi-Williams distribution to process them. They detected the tooth breakage of a gear from the dynamic load profiles.

Demetgül et al. studied the misalignment of pillow blocks. In this study, the characteristics of time series data were extracted using statistical methods. The data of the statistical results were classified and decided using automated machine learning (AutoML), SVM, gradient boosting (GB), and auto-multilayer perceptron (AutoMLP) methods. An attempt has been made to determine the most appropriate algorithm.[36] Demetgul et al. investigated the effectiveness of statistical and short-time Fourier transform (STFT) on pillow block irregularities.[8,9] In another study, Demetgül et al.[37] converted the time series data into images without using feature extraction algorithms and determined the horizontal and vertical axis misalignment of the pillow blocks using the revised Resnet algorithm. The most recent version of these studies is the study of adaptation to real CNC machines. In this research, the axis misalignment in the pillow block is adapted to real CNC machines by using transfer learning. In this study, the performance of time series based deep learning algorithms (LSTM, 1DCNN, AutoEncoder) was compared.[38] Carreon et al.[39] introduce a novel signal segmentation analysis technique for detecting and localizing linear axis guide rail misalignment, aiming to improve the accuracy of misalignment detection and localization in applied sciences.

When all the literature studies and our studies are evaluated. These studies focused on the axis misalignment of the pillow blocks. However, axis misalignment occurs at different locations. This study focuses on a different point. In this study, the axis misalignment between the moving main table and the shaft was investigated.

In many systems, axial misalignment occurs between the table and the spindle over time or during assembly, and its detection takes time. Since CWT images are simple, deep and complex deep learning algorithms encountered overfitting and underfitting and did not perform well. Simple and not deep LeNet5 algorithm performed well. Hyperparameter tuning, transfer learning, and ensemble technique were used to further improve its performance. In Ensemble, the CNN Lenet5 structure and parameters, which were found to be the most suitable with hyperparameter techniques for feature extraction, were used. Instead of the Fully Connected Layer section, the RF algorithm was used for classification and decision making. Thus, the performance of the algorithm was increased to 100%. Additionally, unlike the literature, very small axis misalignments were examined.

In addition, while other studies used time series data, this study converted motor current data into images using CWT. It was found that simple structured deep learning was sufficient in this way.

Theoretical background

In this section, misalignment at the machine tools and additive manufacturing systems will be discussed. Continuous wavelet transformation (CWT), Convolutional Neural Networks (CNN), and LeNEt will be briefly outlined. To enhance the performance of the existing classification methods, ensemble method and hyper parameter tuning is used in this study. These approaches are presented at the rest of the section.

Misalignment

Misalignment occurs when two shafts are not on the same centerline (Figure 1). The alignment between the motors and the shaft of the load needs attention. Although the minor combined misalignment is seen as a normal working condition, when the misalignment increases, the static and dynamic load on the bearings will increase significantly. To avoid this failure, misalignment needs to be monitored during the operation and should be detected when the misalignment is out of tolerance.[8,9,40] In this study, horizontal and vertical axis deviations were created in the experimental set, as shown in Figure 1(b). The friction between the table and the rotor increases in these axis misalignments.

Continuous wavelet transform (CWT)

To identify the misalignment fault at different speeds of the ball screw drive, the signal processing methods in the time domain or the frequency domain are insufficient. Unlike STFT, the scale and time-shifting enable the analysis of the signal with different resolutions at different frequencies. The wavelet transform has a high resolution in the frequency domain at the low
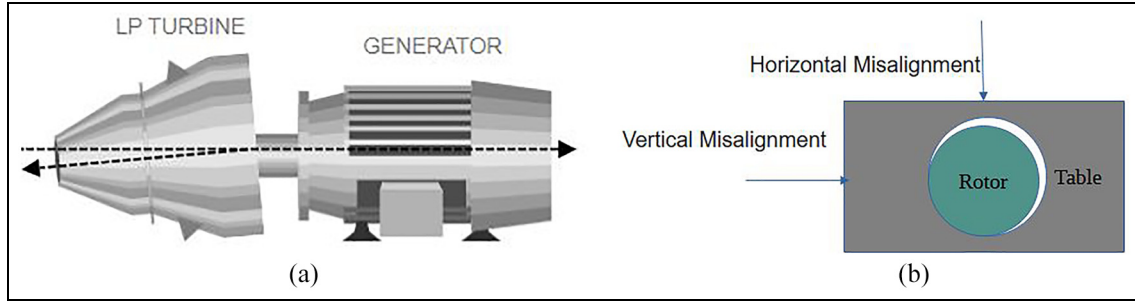
Figure 1. Misalignment at mechanical systems.
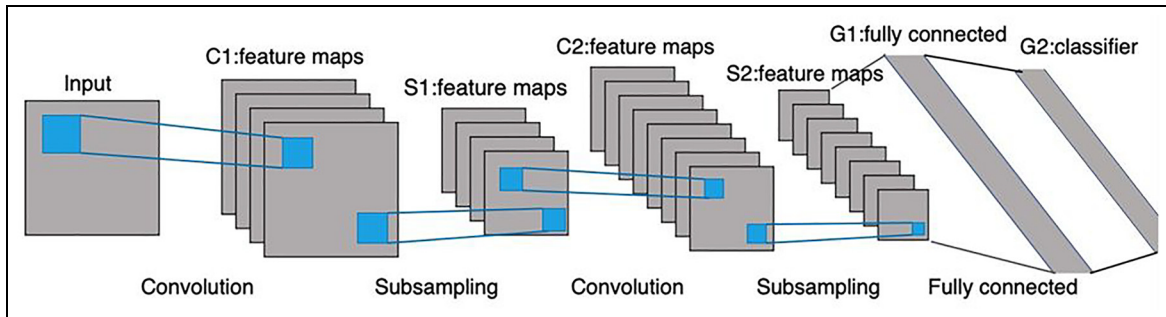
Figure 2. The structure of LeNet.[15]

### Convolution neural network (CNN) and LeNet5

CNN is a supervised deep learning method, which has achieved success in image identification,[42] face recognition,[43] and speech processing.[44] It has shown good performance when it worked with 2D data. The LeNet consists of two convolution layers, two pooling layers, and two fully connected layers as shown in Figure 2. After the operation of convolution layers, the number of feature maps may change depending on the dimension of the kernel. A fully connected neural network is introduced as a classifier in this structure. The input of the classifier is one dimension vector. The dot product between weight vector and input will be calculated. Added by bias, the output will be passed to the activation function for classification. Besides the structure of CNN, there are some hyperparameters to be determined, for example, the number of convolution layers and pooling layers, the size of kernels for each convolution layer, and the size of the pooling. These parameters majorly depend on the size of the input, which is the result of CWT in this study. As mentioned before, Lenet CNN is a supervised learning method, which means it should be trained with labeled data. The idea of training the convolution layers is to adjust the weight in the kernel to minimize the loss function. The loss function is calculated using squared error and back propagated from the fully connected layer using the chain rule. The convolution layers perform the back propagation using a reversed convolution kernel while the pooling layers use the up sample.

The LeNet structure was used for misalignment detection in this paper. The number of units in the last fully-connected layer corresponds to the number of classes. The number of units, feature maps, and the kernel size in other layers are the hyperparameters. The results of continuous wavelet transform were divided into seven classes: normal condition, 0.025, 0.05 mm, 0.075 horizontal misalignments, and 0.05, 0.1 mm 0.15 vertical misalignments. This means the speed and stroke information were not provided for classification although available. The data were labeled with their class. Based on experience, 80% of the data were selected randomly to train the LeNet. The rest 20% of the data were used as test cases to evaluate the performance of the LeNet. The LeNetmodel was built using the Tensorflow library. The default parameters of the LeNet were selected based on the gained experience at the misalignment experiments.

Ensemble learning

Ensemblelearning combines different learning algorithms to obtain better classification performance. Machine learning algorithms have their own advantages and limitations. A single machine learning algorithm may not make the perfect prediction. Combiningof machine learning algorithms eliminates these limitations and obtains better performance.[17, 45, 46]
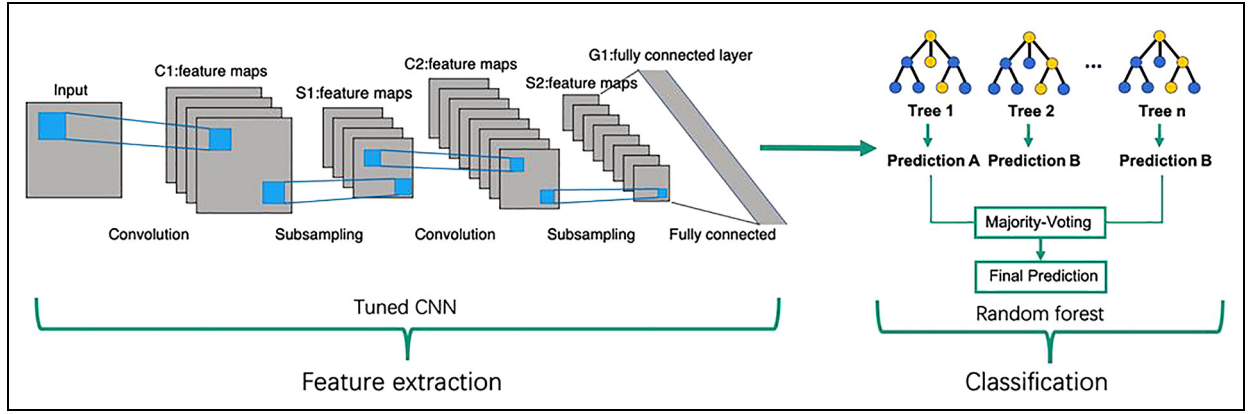
Figure 3. The structure of ensemble method: a combination of CNN and random forest.

There are three major ensemble methods, namely bagging, stacking, and boosting.

Bagging trains different classification models on different samples of the training data set. Stacking uses model learning to combine the predictions of different classifiers. Boosting adds ensemble models that correct the prediction of the previous model sequentially.

Bagging was used in this study. RFclassifiers and combined RF classifiers with the Revised LeNetwere used. The LeNet CNN model extracted the features, and the RF made a prediction on the given data according to the extracted features. The schematic structure of the proposed ensemble method is shown in Figure 3. The RF classifier was connected directly to the output of the fully connected layer of the CNN model.

Hyperparameter tuning

The hyperparameters are the model configuration arguments manually by the developer manually to guide the entire learning process. The effect of the hyperparameters on the CNN model is known. Finding the best combination of hyper-parameters for any given data set is challenging since the hyperparameters tend to interact non-linearly.

One approach to overcome this challenge is to objectively search for different combinations of the hyperparameters for the CNN model and select a subset that gives the best performance. First, the search space and its bounds are defined. The search space can be described as an n-dimensional space, where each hyperparameter represents one dimension. The scale of the dimension is the possible values that the hyperparameter can take. The values can be an integer, a float, a string, etc. So the hyperparameter can be thought of as one point in the search space or a vector.

There are several search methods such as grid search, random search, and Bayesian optimization. Random search was used in this study since it is one of the simplest and most widely used optimization methods. It also has the advantage of obtaining hyperparameter combinations that developers would not intuitively choose, although it may take more computation time. The random search method samples randomly in the search space after defining the bounds of the hyperparameters. After sampling, the training process is run for several epochs and the model is evaluated by using the test data set. The combination of hyperparameters that gives the highest test accuracy is selected

Random search tuning takes a lot of time, especially without a GPU. The step length between two points in the search space cannot be very small. And the training process of CNN is not reproducible. Therefore, the tuning process with random search can optimize the performance of the CNN model.

The random search method requiresto define the search space to be defined as a bounded domain of hyperparameters before sampling randomly sampling in the domain using the KerasTune library. The hyperparameters tuned in the random search are the e number of unitsand the activation function in the convolution layers and fully connected layers. The range of unit numbers was set from 32 to 128 with the increments of 32. The choices of activation functions is ReLU, Sigmoid, SoftMax, and hyperbolic tangent function. The number of units for the last fully connected layer was not tuned as it was determined by the class number (Table 1).

After the tuning process, the best test accuracy was 96.43% and it was achieved with 32 units in the first convolution layer with ReLU, 96 units in the second convolution layer with ReLU, and 64 units in the first fully connected layer with ReLU. The activation

Table 1. Summary of the search space.

Hyperparameters	Search space
Conv2d_1 units	'min value': 32, 'max value' 128, 'step': 32
Activation_1	'ReLU', 'tanh', 'sigmoid'
Conv2d_2 units	'min value': 32, 'max value' 128, 'step': 32
Activation_2	'ReLU', 'tanh', 'sigmoid'
Dense_1 units	'min value': 64, 'max value' 128, 'step': 32
Dense_1 activation	'ReLU', 'tanh', 'softmax'
Dense_2 activation	'ReLU', 'tanh', 'softmax'

Table 2. Summary of the CNN model after hyperparameter tuning and adding an extra layer.

| Layer (type)	LeNet output shape	New output shape	New output shape with adding an extra layer	Number of parameters
Conv2d_1	(57, 498, 32)	(57, 498, 32)	(57, 498, 32)	320
Activation_1 (**ReLU**)	(57, 498, 32)	(57, 498, 32)	(57, 498, 32)	0
Max_pooling2d_1	(28, 249, 32)	(28, 249, 32)	(28, 249, 32)	0
Conv2d_2	(26, 247, 32)	(26, 247, **96**)	(26, 247, **96**)	**27744**
Activation_2 (**ReLU**)	(26, 247, 32)	(26, 247, **96**)	(26, 247, **96**)	0
Max_pooling2d_2	(13, 123, 32)	(13, 123, **96**)	(13, 123, **96**)	0
Conv2d_3			**(24, 245, 32)**	**27680**
Activation_3 (ReLU)			**(12, 122, 32)**	**0**
Flatten	51168	**153504**	46848	0
Dense_1 (ReLU)	64	64	64	**2998336**
Dense_2 (softmax)	7	7	7	455

It shows the bold values selected from the table.

function for the last layer was SoftMax. The model had one convolution layer more than LeNet, with 32 layers. After running the algorithm, the best test accuracy was 97.14%. The summary of the new CNN model is shown in Table 2. The test accuracy improved significantly after the hyperparameter tuning by changing the number of feature maps in the second convolutional layer and the activation function. Selected parameters are marked in bold.

Experimental setup

A linear stage was used in the tests (Figure 4(a)). It was designed to introduce different misalignments in the assembly and to conduct the experiments. The experimental set was developed in kit wbk for this purpose. Experiments were conducted in laboratory environments that were not affected by ambient noise and temperature changes. The linear stage consisted of a ball screw, a servo motor with the controller, bearings with their housing, a table, ball rails, guideways, the nut bracket, and the coupling. The assembly was installed on the pillow blocks. The servo motor was Beckhoff AM8062. It was coupled with the ball screw. The

motor rotor rotations turned the ball screw and created the linear motion of the stage. The table could move between pillow block 1 and pillow block 2 with a maximum stroke of 600 mm. Two limit switches were assembled on each end of the stroke for safety purposes. A screw based misalignment simulation mechanism related to axis misalignment was designed for this study. The controller used the Twincat software. The program moved the workpiece from the starting position to the desired position. The linear stage was designed to be able to set the different levels of misalignment. Figure 4(b) shows that any vertical and horizontal misalignment can be selected by tightening two sets of screws. The screws pushed the ball screw away from its normal working position and an angular misalignment was created. In the normal working condition, both sets of screws are loose and we assume the shafts are perfectly aligned in this situation. The misalignment is measured using the calibrator. To evaluate if the external force on the table can be detected using the current signal, the force will be applied to the table while moving from pillow block 2 to pillow block 1. To generate the force and simulate the working condition with abnormal force on the table, a pneumatic cylinder
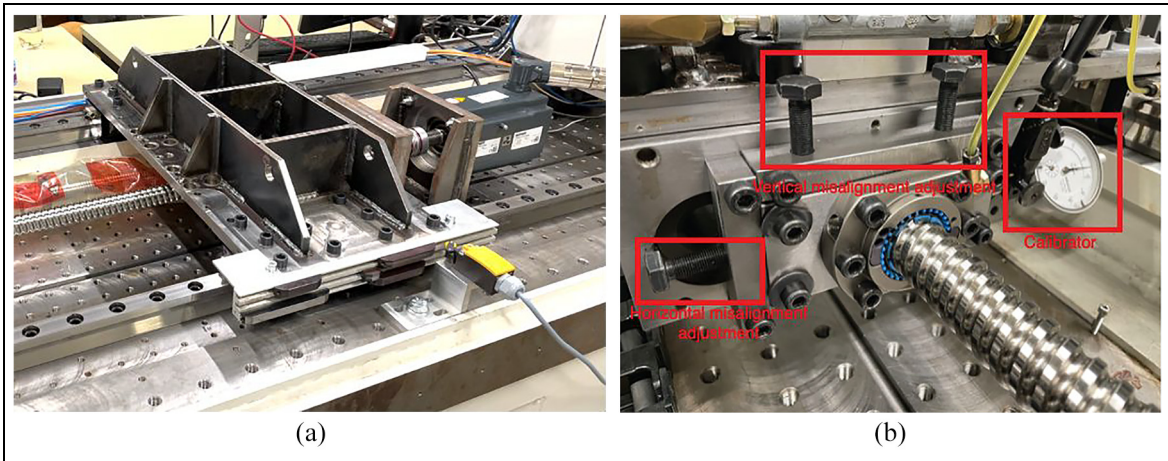
Figure 4. Experiment setup in wbk Institute of production.
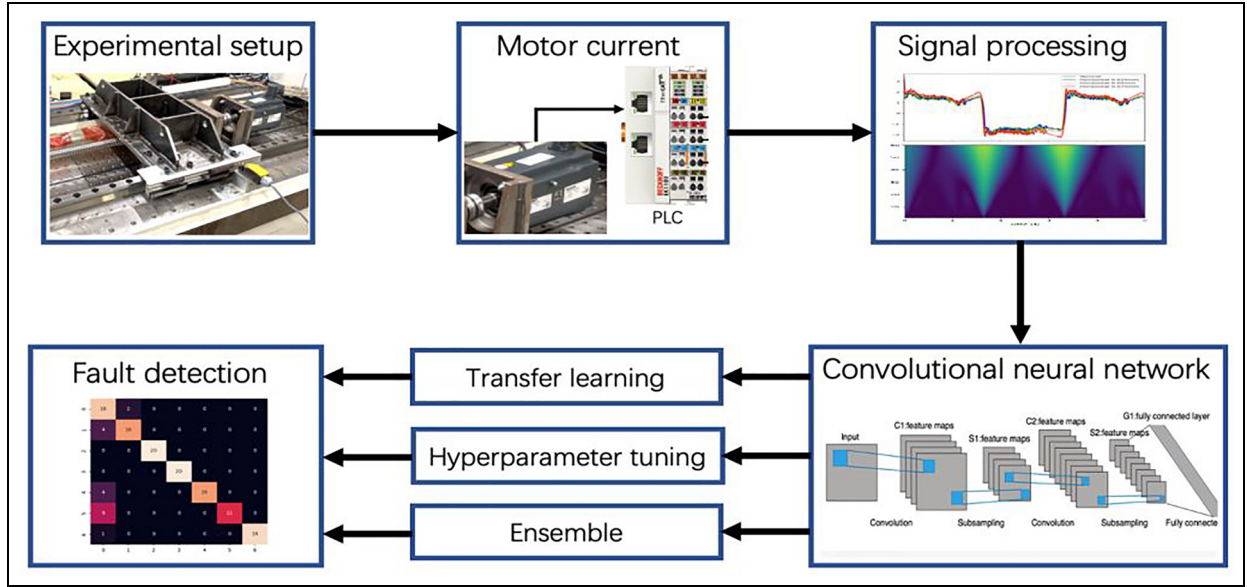
Figure 5. Workflow of the proposed method.

is attached. An aluminum cylinder holder let the piston rod of the cylinder and the table to stay at the same height. So that the resistance from the piston rod will be applied to the table directly. The pressure in the tube is measured with a pressure sensor. The amount of the force applied to the table is controlled by adjusting the regulating valve connected to the cap-end port of the cylinder. The experimental set was run at short intervals and data was collected. There is no working machine around the experimental set. The experimental set is fixed to the floor against vibrations. Therefore, their impact has been minimized. In the study, MITUTOYO 2109S-10 Dial Indicator 0.001 × 1 mm was used to simulate axis misalignment. It was also used to set the normal state. Motor current data was taken according to different screw looseness. The lowest level was taken as the normal state, the indicator was reset and the axis misalignment value was adjusted according to this reference point.

Experimental data collection and working with the CNN model

Implementation of the proposed method is presented in Figure 5. Motor current was monitored while conducting the experiments in the normal working condition and with different levels of vertical and horizontal angular misalignment settings. The motor current signal was recorded from the PLC and saved as a CSV file using the TwinCAT software. No additional sensors were needed.

The current signal was processed with the continuous wavelet transform using the PyWavelets library. The scalogram, which is the CWT's result, was labeled with the alignment condition during the experiment and saved in a list or tensor. The labeled data were the input

images for the CNN model for misalignment diagnosis. The list of labeled misalignment data was separated into training data and test data randomly. After the LeNet CNN model was trained, its performance was evaluated on the test data. Hyperparameter tuning and ensemble methods were employed to improve the performance of the CNN model. It consists of Ensemble part LeNet and RF. Features of graphs created with CWT are extracted with Lenet. Then, RF algorithm is used for classification. The same experiments were repeated after different forces were applied to the table.

The LeNet CNN model trained with different levels of vertical and horizontal misalignment was used as the pre-trained model. With the help of transfer learning theory, a new model for classifying the force variation was trained based on the pre-trained model. The force variations were also randomly selected for training and test data. Since the feature extraction part of LeNet CNN was already trained, the parameters in convolution and pooling layers were set non-trainable during the training, which means they were not changed to minimize the loss function. Only the parameters in fully connected layers were updated during the training. The result of transfer learning was also evaluated with the test data. Using the same pre-trained model, another LeNet CNN model was trained and evaluated for detection of the forces and forces combined with misalignment in the vertical and horizontal axis. A large amount of data was needed to train the LeNet CNN model. Because there was no relative data set available, all of the experiments were carried out in the component laboratory. Each experiment was repeated 25 times. For misalignment detection, 0.025, 0.05, and 0.075 mm of displacement were made for horizontal misalignment and 0.05, 0.1, and 0.15 mm of displacement were introduced for vertical misalignment. The table operated at the speed of 100 and 200 mm/s. The

Table 3. Experiment plan for misalignment.

Condition	Stroke	Speed	Repeat times
Healthy	100 mm	100 mm/s	25
		200 mm/s	25
	200 mm	100 mm/s	25
		200 mm/s	25
Horizontal misalignment 0.025, 0.05, 0.075 mm	100 mm	100 mm/s	75
		200 mm/s	75
	200 mm	100 mm/s	75
		200 mm/s	75
Vertical misalignment 0.05, 0.1, 0.015 mm	100 mm	100 mm/s	75
		200 mm/s	75
	200 mm	100 mm/s	75
		200 mm/s	75
Force variation (10, 25, 50 N)	170 mm	50 mm/s	75
Force variation (25,50 N) + horizantal and vertical misalignment (0.05 mm)	170 mm	50 mm/s	100

stroke was 100and 200 mm. The experiment plan for different misalignment conditions is shown in Table 3. For force variation estimations, the effect of external forces on misalignment detection needed to be evaluated. Therefore, the friction of the pneumatic cylinder was seen as the friction on the guideways under the normal working condition. Since we cannot measure the force directly, pressure was measured. 10, 25, and 40 N forces were applied by the cylinder and an adjustable valve connected to the cap-end port when the table pushes the piston rod back into the barrel. The force applied to the table was also combined with the horizontal and vertical misalignment in the experiments. Experiments were repeated 25 times for training and testing as well. One hundred seventy-five experiments were conducted at different load applications. Totally 875 experiments were performed for the study.

Results

The motor current was obtained from the test bench via PLC. The current value reflected the motor torque. Therefore, the curves of the current signal reached the peaks when the table was accelerated, or brake was applied. The curves were flat when the table moved at a constant speed. Figure 6(a) to (b) show the current signal acquired under normal conditions, horizontal misalignment, and vertical misalignment. The current value increases when the degree of misalignment increases.

Scalogram of CWT

To visualize the output of continuous wavelet transform, the scalogram was plotted using the matplotlib library. Scalograms at the normal conditions and at various horizontal misalignments are presented in Figure 7. Scalograms at the normal and at the various vertical misalignments are presented in Figure 8. In both cases, the data was collected when the linear stage moved forward and backward 4 times with 200 mm/s speed. The stroke distance was 100 mm. At

the scalograms, the horizontal axis had a range of 0 to 5 s on the time axis, and the vertical axis had a range of 0 to 60 in the scale. A higher scale means lower frequency in these figures because the frequency is divided by the scale. The color bars corresponded to the quadratic absolute values. The scalograms showed high time-resolution in high frequency and high frequency-resolution in low frequency. In the figures, the absolute value was greater than 0 when the motor started and braked, which corresponds with the current signal in the previous section. For horizontal misalignment, the absolute value increases when the misalignment increases, even for 0.025 mm misalignment. But for vertical misalignment, the 0.05 mm misalignment is not easy to identify with the human eyes. In some experiments, the maximum magnitude of 0.05 mm misalignment is even lower than the normal condition, which means the load on the motor is smaller than in the normal conditions. This implies that the misalignment in the vertical axis decreased when the screws pushed the nut downwards.

When the table runs at even speed, the magnitude is close to or equal to 0. It can be observed from the figures that the scalogram magnitude is higher when the table is moving back to the initial position than it is moving to the desired position. The color change is more obvious when the misalignment increases. The only exception is a 0.05 mm misalignment in the vertical axis.

Studies to increase the performance of the algorithm

Before adding new convolution layer. The performance on the test data improved significantly after the hyperparameter tuning by changing the number of feature maps in the second convolutional layer and the activation function. Figure 9 shows the training accuracy and test accuracy during the training process of the tuned CNN model. After the tuning process, the best test accuracy in the log is 96.43% and it is achieved with 32 units in the first convolution layer with ReLU, 96 units in the
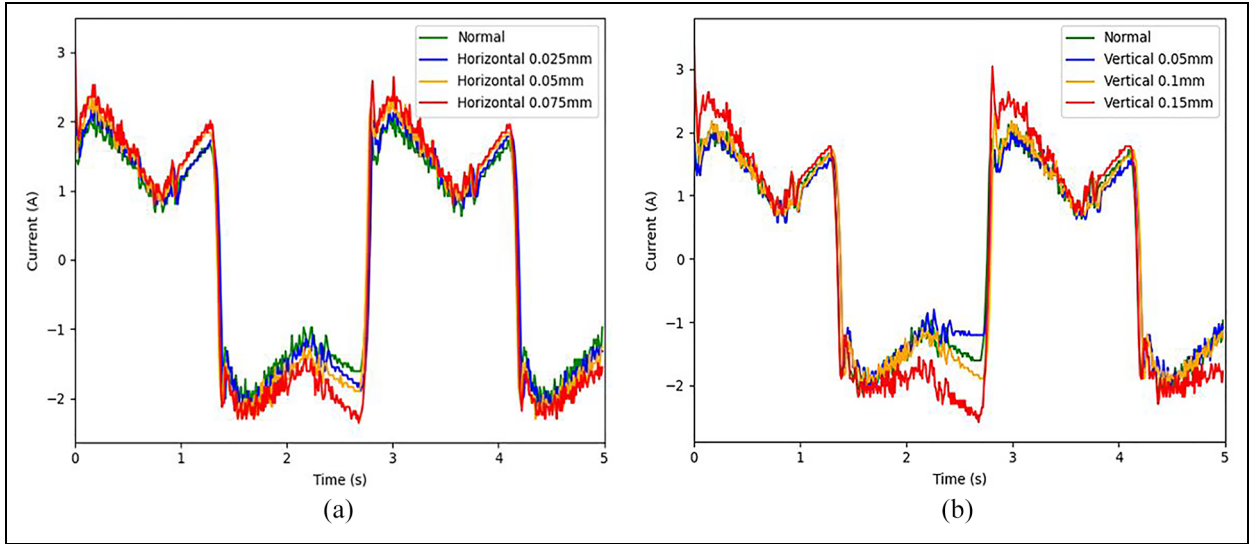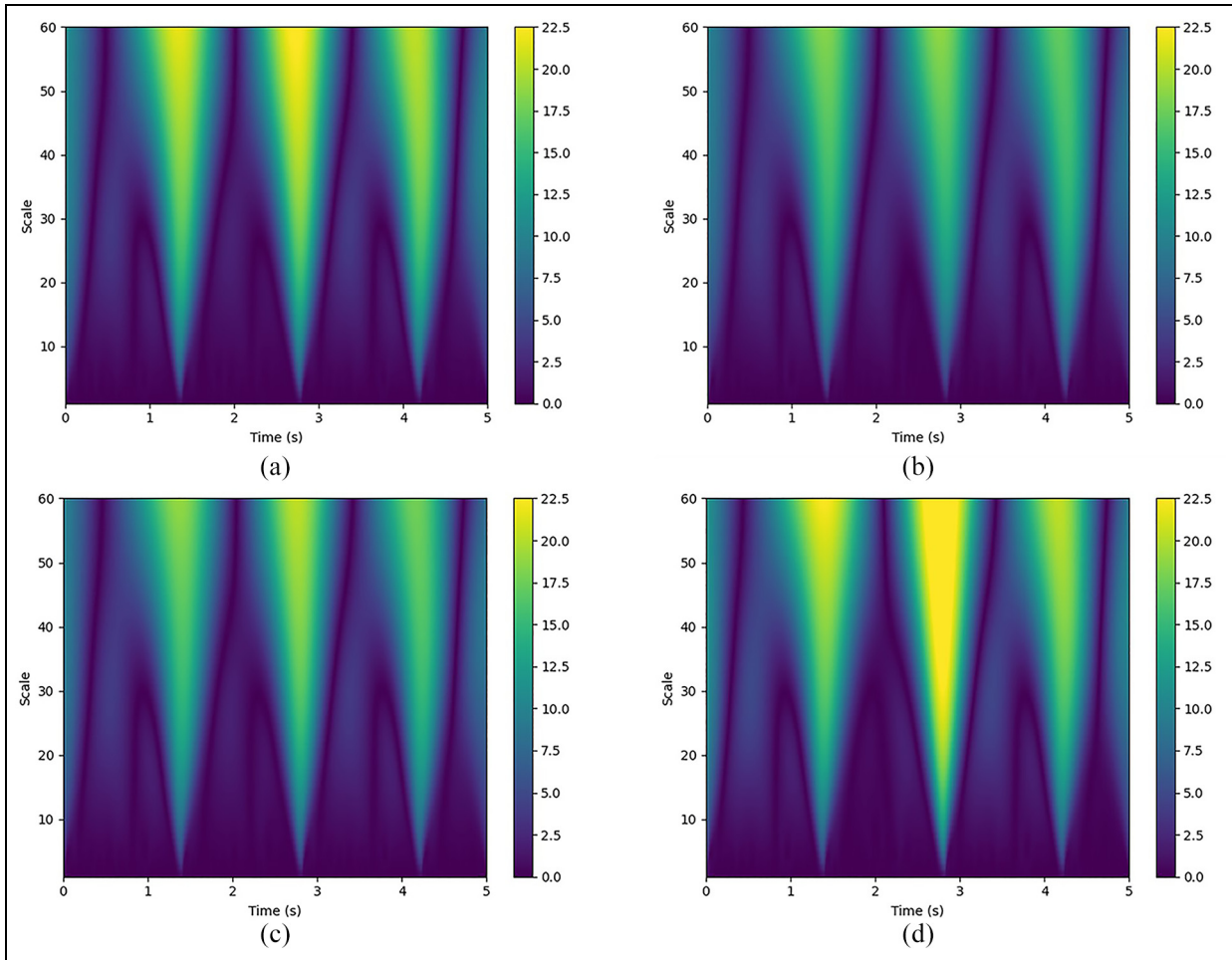
Figure 6. Motor current raw data (a) Horizontal misalignment (b) Vertical misalignment, stroke 100 mm, speed 200 mm/s.
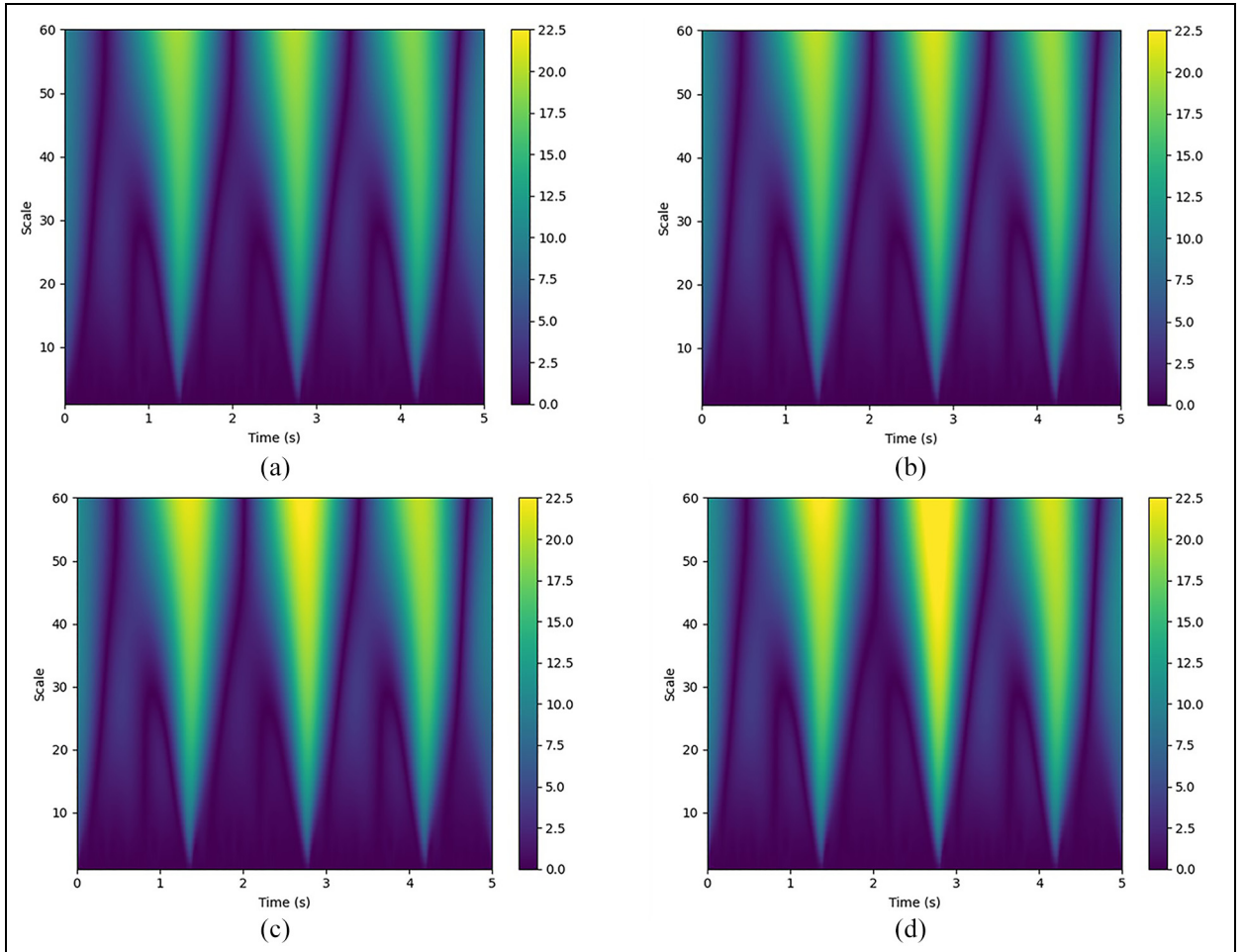

Figure 7. CWT scalogram: horizontal misalignment, stroke 100 mm, speed 200 mm/s: (a) Normal, (b) Horizontal 0.025 mm, (c) Horizontal 0.05 mm, and (d) Horizontal 0.075 mm.

second convolution layer with ReLU, and 64 units in the first fully connected layer with ReLU. The activation function for the last layer is softmax. The orange line is the test accuracy, and the blue one is the training accuracy. Although the performance of the model is improved, the figure shows the overfitting of the model

Figure 8. CWT Scalogram: Vertical misalignment, stroke 100 mm, speed 200 mm/s: (a) Normal, (b) Vertical 0.05 mm, (c) Vertical 0.1 mm, and (d) Vertical 0.15 mm.
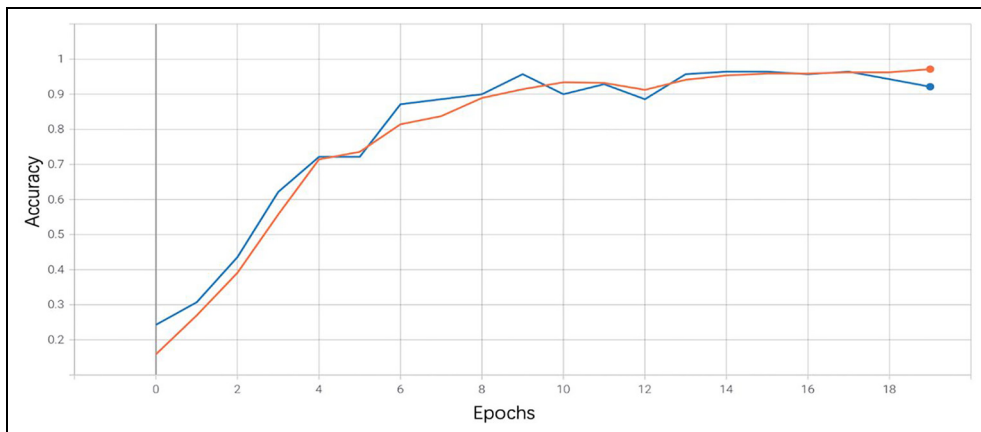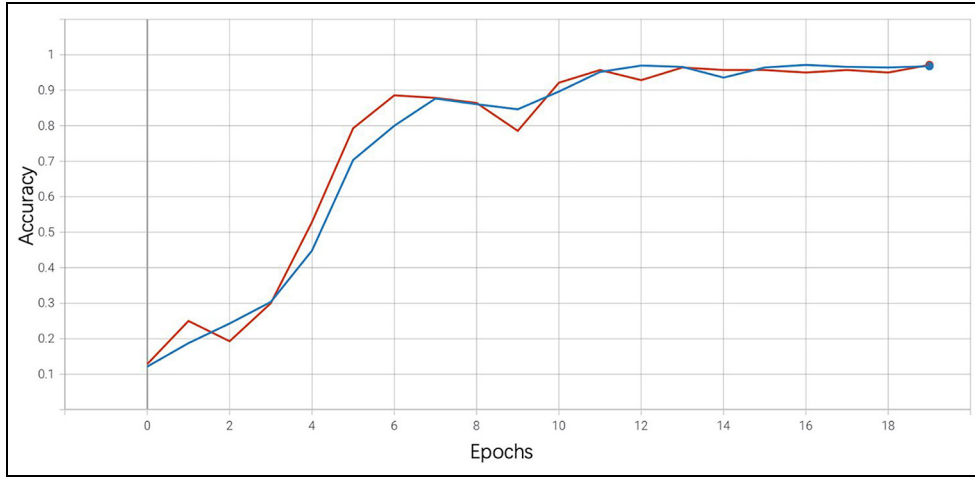

**Figure 9.** Training (blue) and test (orange) accuracy over epoch.

after 17 epochs of training. This training structure can be seen in Table 2 under the new output shape.

After adding new convolution layer. After running the algorithm, the best test accuracy in the log is 98.33%. It is reached with the structure shown in Table 2. The

model has one more convolution layer compared to LeNet. The extra convolution layer also reduces the number of parameters in the model. The variation of LeNet is trained and tested with the same data set as in the search process to reproduce the result and to save the model for further classification. The batch size and epochs remain the same. Figure 10 shows the training

Figure 10. Variation of LeNet: Training (blue) and testing (red) accuracy.
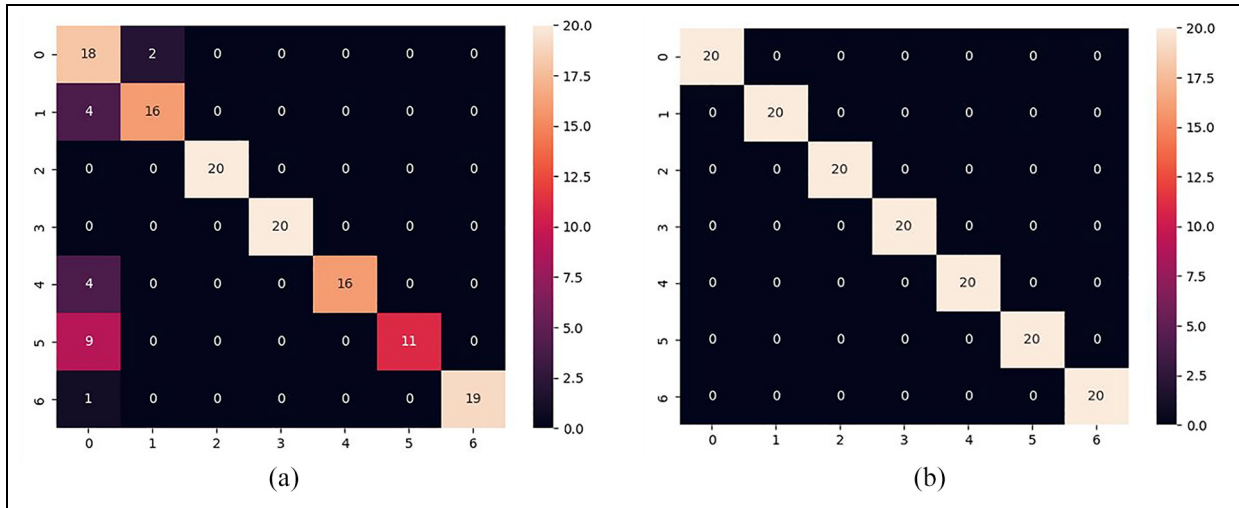
Figure 11. Testing confusion matrix result: (a) LeNet and random forest without hyperparametric tuning and (b) LeNet and random forest with hyperparametric tuning.

and test accuracy over the epoch. The blue line is the training accuracy, and the red line shows the validation accuracy. Compared with the tuned CNN model in the previous section, the training and testing accuracy increase slowly in the first four epochs of training. The test and training accuracy converge after 20 epochs of training. The test accuracy reaches 97.14%, which is higher than the others. There is no overfitting shown in the 20 epochs of training.

Ensemble learning. The ensemble method was used in this study and the tuned convolutional neural network model was combined with a random forest classifier. This combination can be explained as the CNN model extracts the features of different alignment conditions and the random forest makes a prediction on these features. To implement it, the last fully connected layer of the CNN model that makes the prediction was deleted.

The output of CNN's third last layer was passed to the random forest.

In the previous section, the CNN model was trained and tuned with the training data set. The feature extraction part of pre-trained CNN or tuned CNN could be used directly without further training. The training part of this method is to fit the random forest with the training data set. The number of decision trees in the random forest was set to 100. After fitting the random forest, the performance of the ensemble method was evaluated with the test data.

The confusion matrix in Figure 11 shows the result of the ensemble method. Class 0 is the healthy condition. Class 1 to 3 is the 0.025 mm, 0.05 mm, and 0.075 mm misalignment in the horizontal axis. Class 4 to 6 are the 0.05 mm, 0.1 mm, and 0.15 mm misalignment in the vertical axis. As shown in Figure 11, there were 140 test cases. In Figure 11(a), the ensemble method used pre-trained CNN model without tuning

Table 4. The misalignment detection accuracy.

Method	Accuracy
CNN (LeNet)	83.57%
CNN with hyperparameter tuning	96.43%
Variation of LeNet	97.14%
Ensemble (CNN + random forest)	100%

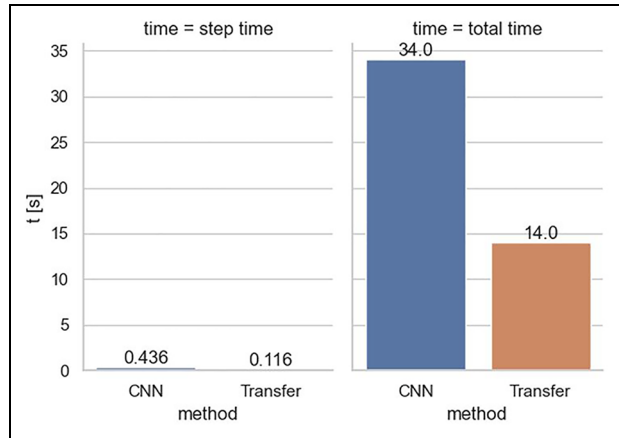**Table 5.** Force and misalignment detection accuracy with transfer learning.

Detection problem	Method	Accuracy
Force variations	Transfer learning	100%
Force + misalignment	Transfer learning	94.29%
Force + misalignment	Transfer learning + batch size optimization	97.14%

Figure 12. Step time and total training time of normal CNN and transfer learning.

misclassifies 9 out of 20 test cases with 0.1 mm vertical misalignment, 4 of 0.025 mm horizontal misalignment, and 0.05 mm vertical misalignment, and 1 of 0.15 mm vertical misalignment to a healthy condition. 2 of healthy condition testing samples were misclassified as 0.025 mm horizontal misalignment. The overall testing accuracy was 85.72%.

For the ensemble of the tuned CNN model and random forest, the test accuracy was 100%. All the test cases were classified correctly as shown in Figure 11(b).

The collected data was processed with continuous wavelet transform and divided into training data and testing data randomly. A CNN model (LeNet) was trained to detect the angular misalignment between the ballscrew and the motor. The performance of the CNN model improved by using hyperparameter tuning, changing the structure of LeNet, and ensemble with a random forest classifier. After that, the impact of changing the ballscrew on the detection accuracy was evaluated. Although the experiment results of different alignment conditions and the anomaly detection accuracy are presented in Table 4. The results show that the proposed motor-current-based misalignment fault detection method is capable to detect different levels of misalignment. Table 4 also shows that with the help of hyperparameter tuning and ensemble, the accuracy was improved significantly.

For different force variations applied to the table, the transfer learning method was used based on the tuned CNN model for misalignment detection. The detection accuracy of different forces was 100%. With 0.05 mm misalignments in vertical or horizontal directions, the detection problem was more complicated. The detection accuracy was 94.29% by using the transfer learning and it reached 97.14% after the training parameters were optimized in the transfer learning. The detection accuracy in Table 5 shows that it is feasible to use the proposed method to monitor the alignment condition even if the table was exposed to different external forces.

The batch size was set to 32 and the number of epochs was set to 20. The results were compared with the transfer learning result with the same training parameters. Both training processes were carried out using the same Quad-Core Intel i5 processor. The time of each step in the training and the total training time is shown in Figure 12. The step time is the time needed to train the model with 32 training examples (1 batch) and to make one gradient update. The total training time is the time needed to train the model with all the training examples for 20 epochs. The step time of training a new CNN model was 3.76 times higher than adapting the pre-trained CNN model with transfer learning. The total training times were reduced significantly by using transfer learning. The accuracy rises faster than training a new CNN model. The overall efficiency of the transfer learning was very good.

Conclusion

In this work, an experimental setup was designed to simulate the different alignment and operating conditions of a linear stage with a ball screw. Normal and three levels of misalignment in the horizontal and vertical directions were tested. Different load conditions were also considered. A sensorless misalignment monitoring system using deep learning was proposed.

The CWT extracted the information from the time domain data to easily estimate the misalignment from the time-frequency domain plots. It also simplified the detection of misalignment even when the table has different stroke lengths and moves at different speeds, because it provides the information in the time and frequency domains simultaneously. Although the stroke length and speed of the table are known, passing them to the classifier complicates the condition monitoring system. The continuous wavelet transform eliminated the need to pass this information.

The misalignment estimation algorithm evaluated the performance of several approaches. The CNN-based LeNet estimated the misalignments with 83.57% accuracy. Hyperparameter tuning improved the estimation accuracy to 96.43%. However, the ensemble method using CNN and Random Forest together achieved 100% accuracy. CNN was used for feature extraction and Random Forest was used for classification. Hyperparameter tuning was a very efficient way to improve the recognition accuracy. Using these algorithms and the communication protocol between the TwinCAT PLC software and Python, an online alignment condition monitoring system with user interface can be easily built. The proposed motor current based misalignment estimation method is very promising for industrial applications.

In this study, it was found that pillow block axis misalignment and table axis misalignment produced different data. Unlike our previous studies, table axis misalignment was investigated. It was concluded that these two axis misalignments could be detected simultaneously. The transfer learning method allowed the developed deep learning algorithms to efficiently learn to make estimates when different forces were applied to the table. Transfer learning is very useful in industrial settings where it is not possible to run experiments with a combination of different parameters.

The main contribution of this study is the demonstration of the ensemble method to improve the accuracy and the use of transfer learning to include different parameters in the identification process in a very efficient way. Both of these approaches are expected to find good acceptance in industry. In addition, the simplest deep learning method, Lenet, was found to be sufficient.

The sample sizes used in this study were chosen to be appropriate for academic presentation and to systematically outline the proposed method. Due to the nature of the data, deeper network structures were deemed impractical, as they often lead to overfitting. However, in industrial applications, deeper CNN structures with larger datasets might be more applicable. These could involve more layers or multiple residual blocks, trained on extensive data sets where concerns about academic representation do not apply.

Future work could also investigate the feasibility of detecting other machine faults, such as bearing and gear faults, using the proposed method. Additionally, applying the method to a CNC machine could be explored.

Ethical Statement

No such permission was required in this study.

Informed consent/ Patient consent

No such permission was required in this study.

ORCID iD

Mustafa Demetgul https://orcid.org/0000-0001-9385-9305

Data availability statement

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

Reference

1. Altintas Y, Verl A, Brecher C, et al. Machine tool feed drives. *CIRP Annals* 2011; 60(2): 779–796.
2. Butler Q, Ziada Y, Stephenson D, et al. Condition monitoring of machine tool feed drives: a review. *J Manufact Sci Eng* 2022; 144(10): 100802.
3. Brecher C and Weck M. Feed axes in machine tools. In *Machine tools production systems* (vol. 3, pp. 7–106). Springer, Wiesbaden, 2022.
4. Lin JL, Liu JY, Li CW, et al. Motor shaft misalignment detection using multiscale entropy with wavelet denoising. *Expert Sys Appl* 2010; 37(10): 7200–7204.
5. Reddy MC and Sekhar AS. Detection and monitoring of coupling misalignment in rotors using torque measurements. *Measurement* 2015; 61: 111–122.
6. Mohanty AR and Fatima S. Shaft misalignment detection by thermal imaging of support bearings. *IFAC-PapersOnLine* 2015; 48(21): 554–559.
7. Caso E, Fernandez-del-Rincon A, Garcia P, et al. Monitoring of misalignment in low speed geared shafts with acoustic emission sensors. *Appl Acoust* 2020; 159: 107092.
8. Demetgül M, Gu M, Hillenbrand J, et al. Misalignment detection on linear feed axis with FFT and statistical analysis using motor current. *J Mach Eng* 2022; 22(2): 31–42.
9. Mustafa D, Yicheng Z, Minjie G, et al. Motor current based misalignment diagnosis on linear axes with short-time Fourier transform (STFT). *Procedia CIRP* 2022; 106: 239–243.
10. Zheng J, Pan H and Cheng J. Rolling bearing fault detection and diagnosis based on composite multiscale fuzzy entropy and ensemble support vector machines. *Mech Sys Sig Proc* 2017; 85: 746–759.
11. Lei Y, He Z and Zi Y. A combination of WKNN to fault diagnosis of rolling element bearings. *J Vib Acoust* 2009; 131(6): 064502.

12. Wen L, Li X, Gao L, et al. A new convolutional neural network-based data-driven fault diagnosis method. *IEEE Transact Ind Elect* 2017; 65(7): 5990–5998.

13. Lei Y, Yang B, Jiang X, et al. Applications of machine learning to machine fault diagnosis: a review and roadmap. *Mech Sys Sig Proc* 2020; 138: 106587.

14. Tuba E, Bačanin N, Strumberger I, et al. Convolutional neural networks hyperparameters tuning. In *Artificial intelligence: theory and applications* 2021 (pp. 65–84). Cham: Springer International Publishing.

15. Zhu Y, Li G, Wang R, et al. Intelligent fault diagnosis of hydraulic piston pump combining improved LeNet-5 and PSO hyperparameter optimization. *Appl Acoust* 2021; 183: 108336.

16. Xu G, Liu M, Jiang Z, et al. Bearing fault diagnosis method based on deep convolutional neural network and random forest ensemble learning. *Sensors* 2019; 19(5): 1088.

17. Xu Y, Li Z, Wang S, et al. A hybrid deep-learning model for fault diagnosis of rolling bearings. *Measurement* 2021; 169: 108502.

18. Feurer M and Hutter F. Hyperparameter optimization. In *Automated machine learning* (pp. 3–33). Springer, Cham, 2019.

19. Zhang Y, Liu W, Wang X, et al. A novel hierarchical hyper-parameter search algorithm based on greedy strategy for wind turbine fault diagnosis. *Expert Sys Appl* 2022; 202: 117473.

20. Yang B, Lei Y, Jia F, et al. An intelligent fault diagnosis approach based on transfer learning from laboratory bearings to locomotive bearings. *Mech Sys Signal Process* 2019; 122: 692–706.

21. Guo L, Lei Y, Xing S, et al. Deep convolutional transfer learning network: a new method for intelligent fault diagnosis of machines with unlabeled data. *IEEE Transact Ind Elect* 2018; 66(9): 7316–7325.

22. Li X, Hu Y, Li M, et al. Fault diagnostics between different type of components: a transfer learning approach. *Appl Soft Comp* 2020; 86: 105950.

23. Li W, Huang R, Li J, et al. A perspective survey on deep transfer learning for fault diagnosis in industrial scenarios: theories, applications and challenges. *Mech Sys Sig Process* 2022; 167: 108487.

24. Wu Z, Jiang H, Zhao K, et al. An adaptive deep transfer learning method for bearing fault diagnosis. *Measurement* 2020; 151: 107227.

25. Lei P, Shen C, Wang D, et al. A new transferable bearing fault diagnosis method with adaptive manifold probability distribution under different working conditions. *Measurement* 2021; 173: 108565.

26. Li X, Zhang W, Li X, et al. Partial domain adaptation in remaining useful life prediction with incomplete target data. *IEEE/ASME Transact Mechat* 2023; 29(3): 1903–1913.

27. Chen X, Li X, Yu S, et al. Dynamic vision enabled contactless cross-domain machine fault diagnosis with neuromorphic computing. *IEEE/CAA J Autom Sin* 2024; 11(3): 788–790.

28. Guan Z, Chen P, Zhang X, et al. Vibration analysis of shaft misalignment and diagnosis method of structure faults for rotating machinery. *Int J Performab Eng* 2017; 13(4): 337.

29. Kumar A and Kumar R. Development of LDA based indicator for the detection of unbalance and misalignment at different shaft speeds. *Exp Techn* 2020; 44(2): 217–229.

30. Wang NF, Jiang DX and Han T. Misalignment fault detection in dual-rotor system based on time frequency techniques. In *Society for machinery failure prevention technology annual conference 2017: 50 years of failure prevention technology innovation*, MFPT 2017.

31. Jiang Y, Yin S, Dong J, et al. A review on soft sensors for monitoring, control, and optimization of industrial processes. *IEEE Sens J* 2020; 21(11): 12868–12881.

32. Kia SH, Henao H and Capolino GA. A high-resolution frequency estimation method for three-phase induction machine fault detection. *IEEE Transact Ind Elect* 2007; 54(4): 2305–2314.

33. Plapper V and Weck M. Sensorless machine tool condition monitoring based on open NCs. In *Proceedings 2001 ICRA. IEEE international conference on robotics and automation* (Cat. No. 01CH37164) 2001 (Vol. 3, pp. 3104–3108). IEEE.

34. Verl A, Heisel U, Walther M, et al. Sensorless automated condition monitoring for the control of the predictive maintenance of machine tools. *CIRP Ann* 2009; 58(1): 375–378.

35. Putz M, Trimborn C, Naumann C, et al. Sensorless fault detection in linear axes with dynamic load profiles. *Proc Manufact* 2018; 19: 66–73.

36. Demetgul M, Wang W, Fleischer J, et al. AI-based inspection of the axes of machine tools. *Int J Adv Manufact Technol* 2024; 130(5): 2329–2342.

37. Demetgul M, Zihan M, Heider I, et al. Misalignment detection on linear feed axis using sensorless motor current signals. *Int J Adv Manufact Technol* 2023; 126(5): 2677–2691.

38. Demetgul M, Zheng Q, Tansel IN, et al. Monitoring the misalignment of machine tools with autoencoders after they are trained with transfer learning data. *Int J Adv Manufact Technol* 2023; 128(7–8): 3357–3373.

39. Hurtado Carreon A, DePaiva JM and Veldhuis SC. Linear axis guide rail misalignment detection and localization using a novel signal segmentation analysis technique. *Appl Sci* 2024; 14(6): 2593.

40. Chacon JL, Andicoberry EA, Kappatos V, et al. Shaft angular misalignment detection using acoustic emission. *Appl Acoust* 2014; 85: 12–22.

41. León FP and Jäkel H. Signale und systeme. In *Signale und systeme*. De Gruyter Oldenbourg; 2019.

42. LeCun Y and Bengio Y. Convolutional networks for images, speech, and time series. *Handbook Brain Theory Neural Networks* 1995; 3361(10): 1995.

43. Lawrence S, Giles CL, Tsoi AC, et al. Face recognition: a convolutional neural-network approach. *IEEE Transact Neural Networks* 1997; 8(1): 98–113.

44. Abdel-Hamid O, Deng L and Yu D. Exploring convolutional neural network structures and optimization techniques for speech recognition. In *Interspeech 2013* (Vol. 2013, pp. 1173–1175).

45. Zhong SS, Fu S and Lin L. A novel gas turbine fault diagnosis method based on transfer learning with CNN. *Measurement* 2019; 137: 435–453.

46. Toğaçar M, Ergen B and Cömert Z. Classification of flower species by using features extracted from the intersection of feature selection methods in convolutional neural network models. *Measurement* 2020; 158: 107703.