Smart Drill Machines Using Machine Learning

Tianyi Wang
Computer Science
Washington State University

***Abstract*— This project aims to use machine learning techniques to detect anomalies in machine drilling data, classify the data, and eventually improve machining. In this project, we will use thrust force and torque data of the original drill holes. We try to apply a supervised learning on machines and let machines forecast whether the holes have defects or not by analyzing the thrust force and torque data. The ME student will use scanning electron microscope (SEM) to verify the accuracy of the machine predictions. If there is a defect, the ME student will tell us what kind of defect this hole has. In this way, machine will be able to forecast multiple kinds of defects of holes. Our final goal is applying these machine learning frameworks to the real-world drilling equipment or computer numerical control machines.**

I. INTRODUCTION

"Carbon fiber-reinforced plastic (CFRP) composites are widely used in aerospace engineering due to their outstanding mechanical properties, including light-weight, a very high strength-to-weight ratio, a high modulus-to-weight ratio, high resistance to fatigue and good corrosion resistance. However,CFRP have fiber-resin interface on the microscopic level[1]". Besides, compared with other materials, CFRP is really expensive.

Because of the complex composition of CFRP, drilling a good hole on CFRP become difficult and complicated. So how to reduce the error rate when processing CFRP has become the focus of this experiment. In industrial processing engineering, the magnitude of thrust force, torque, the type of drill bit and feeding rate etc. are the most significant factors which can affect error rate during CFRP drilling process. The structure of CFRP has been shown in Fig. 1.

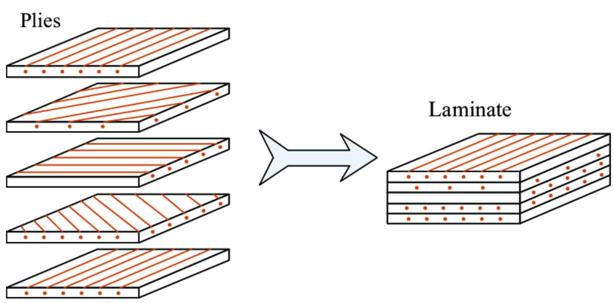


Fig. 1: Structure of CFRP laminates[2]

The target for us is to improve the whole drilling process, which make the whole process generate more good holes. So, the first step for us is how to let drilling machine differentiate which one is good and which one is bad.

On the drilling machine, there are two force sensors. One sensor can measure the thrust force of the drilling head and the other one can measure the torque of the drilling head. If we can build the relationship between roughness coefficients and force curve, then we will be able to let drilling machine to know the quality of the holes which is roughness coefficients. From Doctor Dave, R_a, R_v and R_z can more accurately represent the quality of the holes.

Intuitively, roughness coefficients can represent the quality of the holes and roughness coefficient and quality are directly proportional. The bigger the roughness coefficient are, the worse the hole quality will be. Hence, we can tell the quality of holes according to roughness coefficients. However, in the drilling process, we are unable to know the roughness coefficients. Because roughness coefficients are approached by measuring one hole vertically and on a specific direction after an entire drilling process. The only thing we can got in the whole drilling process is force of two type and time curve.

Some holes have R_a, R_q and R_z. And Some holes have R_a, R_q, R_z, R_p, R_v, R_{sk}, $R_{k\mu}$, R_t. Hence, in this paper, I want to build the relationship between R_a, R_z and force time curve.

The following is the roughness curve of the hole:

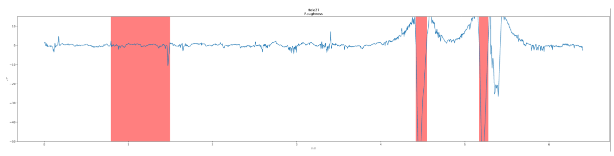


Fig. 2: Roughness curve

The characteristics of a curve can be represented by its troughs and flat areas. The red color represents the trough, and the green color represents the smooth region. In a sense that the total amount of troughs area can represent the quality of current drilling hole.

We try to find the three deepest troughs and all the flat areas in the roughness curve of the hole. If the length of one area is greater than 0.2445mm, max-min<1.899, this area is counted as a flat area, and the depth of the hole is 7.6mm. The speed of the drill bit is 7.62mm. According to this correspondence, we can find the corresponding position on the force curve.

Most of the trough regions correspond to stage 2 on the force data. The upper graph is the corresponding region of the trough in the thrust force and time and torque and time

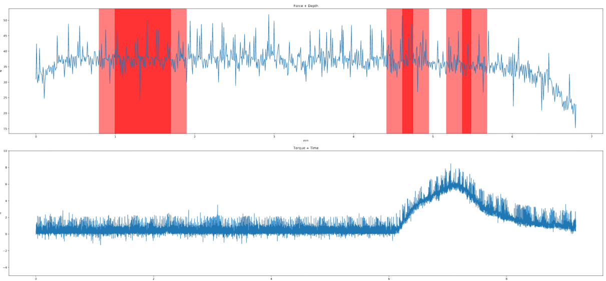


Fig. 3: force + Depth vs torque + time

graph.

Machine learning is one kind of artificial intelligence methods to solve this problem. Machine learning has three subsets, which are supervised learning, unsupervised learning and reinforced learning. This paper are using supervised learning. Hence, we have two directions, which are classification and regression respectively.

- Classification: how to classify all the holes into bad holes set and normal holes set
- Regression: how to forecast the value of R_a and R_z from two kinds of force time curves

Traditionally, we will give rules and data to the machine. Machine will give the result back to us depending on the rules that we have provided. In this experiment, it is subtle that what kind of rules is capable to tell the difference between roughness coefficients and force time curves. The only things we knew are the data from 140 holes. So far, machine learning is able to find the hidden data structure by doing data analysis which is able to help us to solve this problem.

II. PREPROCESSING

A. Staging

When drilling machine is drilling holes on the fiber board, there are five stages in this whole process.

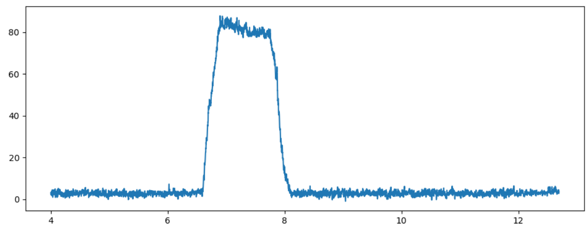


Fig. 4: Thrust force and time

1. The drilling machine does not touch the CFRP board.
2. The drilling head start drilling the board until the drilling head has been totally immersed in the board.
3. The drilling head totally immersed in the board.
4. The drilling head has been partially out of CFRP board.
5. Totally out.

The drilling head has been totally immersed in the board in the second stage. Thrust force and torque are applied to the entire drilling head. Hence the stage 2 is the most important region for us to find relationship between machine status and roughness coefficients. Choosing a proper staging algorithm has been a crucial problem in our research.

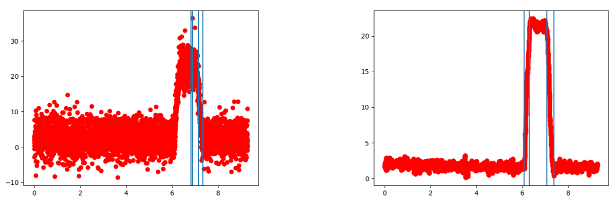
a) Ramer-Douglas-Peucker algorithm

"The purpose of the algorithm is, given a curve composed of line segments, to find a similar curve with fewer points. The algorithm defines 'dissimilar' based on the maximum distance between the original curve and the simplified curve. The simplified curve will consists of a subset of the points that defined the original curve"

This algorithm will divides the line recursively. At first, it will mark the start and the end point. Then it will find the farthest point A from the line segments with the first and last points as end points; Then I will use this point A as starting point and ending point to draw another line segments and do the same operations for the rest point. We have set a threshold k for this algorithm. If the distance between the farthest point and simplified line segment is closer than the threshold k, then this point will be discard. RDP algorithm will stop dividing the line recursively until all the farthest points which are closer than the threshold k have been discarded.

After setting a proper threshold, we can shorten all the points in the stage 1 and stage 3 to two points respectively. After that, we can find the ends of stage 1, stage 2 and stage 3 via finding jumping points.

But, the problem in this method will show up in the situation like hole 22. When the force in the same stage has enormous change, the staging algorithm is unable to have a good performance, The same problem will show up in the following methods too. Hence, the improvement for us is to add Gaussian filter before doing the RDP staging.



(a) RDP

(b) RDP + Gaussian Filter

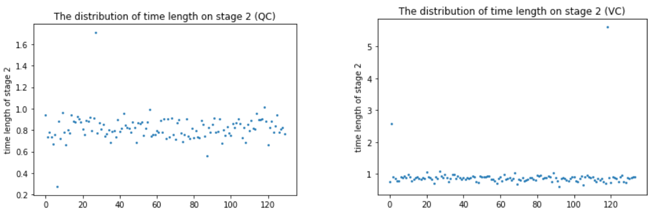
Fig. 5: Segmenting on force-time curve of QC hole 22

b) Naive Windowing

Add Gaussian filter at first, then when we are observing the shape of thrust force-time curve, we will find out stage 1 and stage 3 are really steep. Then we can use a fixed size of window to shift on the thrust force and time curve. The size of window should greater than half the average length of stage 1 and smaller than the average length of whole stage 1. We will do the windowing operation on the beginning of this whole curve. If we can find the jumping points, in another word, the position where the gap between the two ends of the window is the largest. Then we can use this approach to find the start point of stage 1. We mark the position of phase one as $pos1_0$, after $pos1_0$, we will continue the window scan, and then we can find the start position of stage 3, we can mark the start position of stage 3 as $pos3_0$, $pos3_0 == pos2_1$. In the other words, we can find the ending position of stage 2 with windowing approach. In the same way, we can use reverse windowing approach to find the beginning position of stage 2 which is $pos2_0$.

Then the most important question in this problem is how to get the width of the windows. Here I will use Ramer-Douglas-Peucker (RDP) algorithm to do the segmenting at first. Then I will get staging results of each hole, but the staging result is not accurate, we need to do some improvement on this trail. From the staging results, I will make statistics on the length of stage 2:

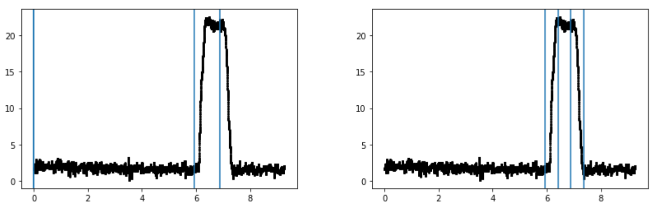
For QC and VC, we browse the whole QC RDP folder and plot the time range length of stage 2 and get the distribution of stage 2 in Fig 7.



(a) QC (b) VC

Fig. 6: The distribution of length Stage 2

Here I sort the whole time length set and get rid of first 3 points and last 3 point of all the points set, which will remove the extreme value. Then we average over this set. The window size should be 0.6 times the average value. The length of any point in the collection should be greater than the width of the window size. This whole process are shown in Fig 8.



(a) First step (b) Second step

Fig. 7: First two steps of naive windowing

This methods is way better than just using RDP algorithm.

B. Fast Fourier Transform

To explore the relationship between different force and time curves and the quality of various holes, the transformation from time domain to frequency domain matters.

DFT stands for transforming discrete-time signals into frequency-domain complex signals. The analysis formulas for discrete-time discrete-time Fourier transforms of finite length and discrete finite length are collectively referred to as discrete Fourier transforms. To keep the natural signal as a digital discrete signal, the first thing to do is to sample the signal. In general, the sampling methods used are equally spaced sampling, that is, the sampling interval between the two samples is the same. Because the process of DFT will consume plenty of time. Then we can use FFT to accelerate this progress.

Currently, we want to build the relationship between thrust force curve and different roughness data. At first, after observing the force time curve, we will find out that force time curve is a segment of vibration changing through time. From Fourier theorem, any periodic function can be regards as a superposition of sine waves of different amplitudes and phases.

Hence, we can understand that any curve that changes with time can be regards as a superposition of different sine and cosine curves. The Fourier formula can help us to transform graph on the time zone into on the frequency zone. Sometimes, we cannot find the relationship on the time zone, we can change our direction into frequency zone via Fourier transform.

After we applying FFT on stage 2 of thrust force and time curve, we will find that the FFT result of force and time curve has too many items which the amount is over 400. The more items, the more detailed the description of the curve. However, if we use machine learning methods on those data, it also will generate more parameters, which will make the whole progress too difficult. If a machine learning method is too complicated, then it will have a better chance to fit the data in the training set better, so which will be more likely to cause over-fitting.

III. MACHINE LEARNING

A. Linear Regression

Linear Regression is one of the methods in supervised learning. In this method, it try to fit all the data point in training set into a straight line as prediction line.

Linear regression is a simple machine learning approach. In the other words, LR has a better effect on reducing the complexity of the model. Linear regression will find the best possible straight line for our data. It will use the difference between prediction and actual value to update itself until it get the best straight line for our data.

Linear regression can be used with FFT data and do prediction on R_a which is one of important roughness coefficients. Here we cut our FFT data into two parts. One parts will be denoted by L (low frequency region), while the other part will be expressed by H (high frequency region). Using the ratio between the sum of L and sum of L + H to predict the value of R_a has been showed in Fig 9.

a) R Area

Still, FFT result for our force time curve is too complicated. Then we need find methods to reduce the complexity. we want to find the relationship between $\{R_a$, R_v and $R_z\}$ and $\{hole\ number,\ r\ area,\ FFT\ data\}$

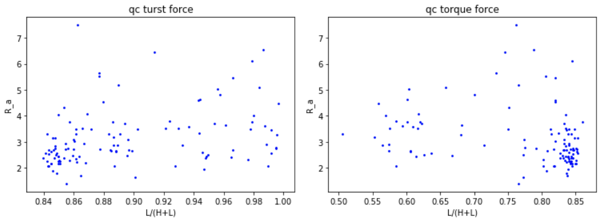


Fig. 8: The prediction on R_a by L/(H + L) among 140 QC holes

If holes of similar quality have low distance between them and holes of different quality have large distance between them. Then this measure can help us classify new holes based on FFT.

On FFT curve, different peaks represent different features. In another word, peaks is able to represent difference of one hole. We want to see if the FFT peaks of different holes but the same quality holes overlap. From another angle, we can understand it as finding similarity between different holes but the same quality. Ordinary least squares method give us an approach to use one straight line to present the whole curve. If the distances between different ordinary least squares straight lines of different holes are close enough, we can think of multiple peaks overlaps on these two curves. Because the peaks are the most important characters of one curve.

From this angle, we can use FFT and linear regression to do the prediction on $R_a\ avg$, $R_v\ avg$ and $R_z\ avg$. From the graph above, we can draw a conclusion which the shape of blue line and yellow line have big difference in r_2, r_4 and r_6 areas. Hence, we can use $\{r_2$, r_4 and $r_6\}$ to do the prediction on roughness coefficients.

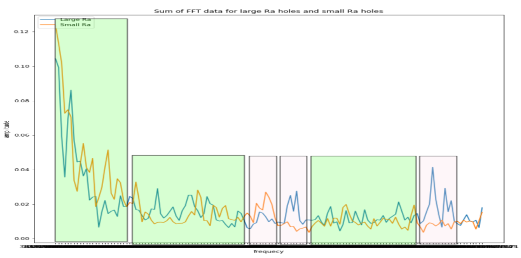


Fig. 9: R area

b) Principal Component Analysis

To reduce the model complexity, we can use feature selection and feature extraction approaches. Feature selection will keep the original data, whereas feature extraction will create brand new one set. PCA is one kind of feature extraction approaches and it is able to reduce the dimension of data.

In the PCA, it will standardize the data at first, then build the covariance matrix. After getting the eigenvalues and eigenvectors of the covariance matrix, it will sort the eigenvalues in decreasing order which rank the eigenvectors.

When we are using FFT data to do the prediction, we will find out there are too many parameters in linear regression method, which will create over-fitting problems. Then we can use feather selection method before doing the prediction which is PCA. We get better result in 20 components PCA of force FFT data and torque FFT data.

B. Classification

a) Mel Frequency Cepstral Coefficents[3]

We can consider the task of finding bad holes as finding a specific set of sounds. Then we can turn the task of classification to speech recognition. In MFCC, we will try to obtain the cepstrum of the frequency spectrum. Because the low frequency component of the cepstrum is the envelope of the frequency spectrum, and the high-frequency component of the cepstrum is the details of the frequency spectrum. These are speech physical information that has been scientifically validated in speech recognition. Here is the spectrum we have obtained from MFCC:

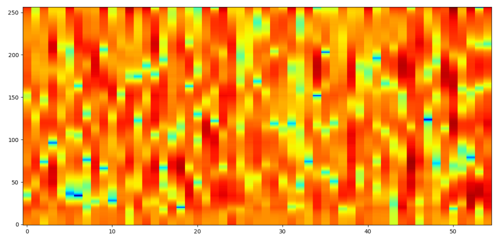


Fig. 10: MFCC spectrum on QC hole 20

b) LeNet-5

Yann[4] et al. proposed a neural network architecture for handwritten and machine-printed character recognition in 1990's which they called LeNet-5. The architecture is a straightforward classification method. In here we use LeNet-5 to do regression on the matrix which is generated from MFCC. Because the LeNet-5 is a method to do the classification, we can use one-hot encoding method in this method, which can transform classification method into a regression method. One-hot encoding can let model learn directly from categorical data with no data transform required.

IV. Evaluation

A. Correlation Coefficients

Till now, we have three methods to do the evaluation: Mean Average Error (MAE), Pearsons and Spearmans. Pearsons and Spearmans are correlation coefficients. Correlation coefficient is a special kind of covariance.

It can reflect if two variables change is in the same direction or in the opposite direction. If same direction, correlation coefficient will be greater than 0, otherwise it will be less than 0. Besides, it is standardized covariance, it is able to eliminates the influence of the changes in the two variables, but simply reflects the similarity of the figure of two variables. Hence, compared to MAE, Pearsons and Spearmans coefficients are more convincing.

```
Model: "sequential_1"
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
conv2d_1 (Conv2D)            (None, 53, 255, 6)        60
_________________________________________________________________
average_pooling2d_1 (Average (None, 26, 127, 6)        0
_________________________________________________________________
conv2d_2 (Conv2D)            (None, 24, 125, 16)       880
_________________________________________________________________
average_pooling2d_2 (Average (None, 12, 62, 16)        0
_________________________________________________________________
flatten_1 (Flatten)          (None, 11904)             0
_________________________________________________________________
dense_1 (Dense)              (None, 4096)              48762880
_________________________________________________________________
dense_2 (Dense)              (None, 512)               2097664
_________________________________________________________________
dense_3 (Dense)              (None, 128)               65664
_________________________________________________________________
dense_4 (Dense)              (None, 16)                2064
_________________________________________________________________
dense_5 (Dense)              (None, 1)                 17
=================================================================
Total params: 50,929,229
Trainable params: 50,929,229
Non-trainable params: 0
```

Fig. 11: The structure of LeNet-5

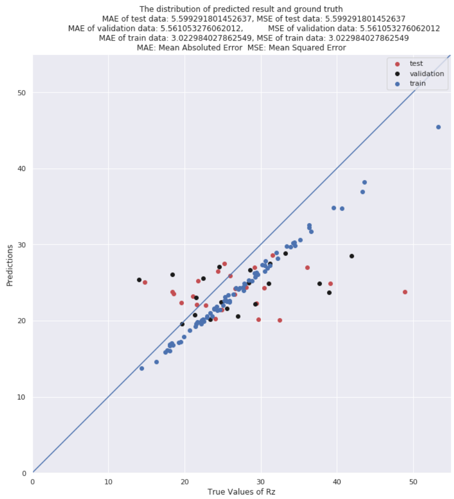


Fig. 12: The prediction situation of LeNet-5

B. Relationship between Hole Number and Tool ware

Using *hole number* with linear regression method to do the prediction has really high accuracy, then we can use Flank wear land (FWL) and Tool wear area (TWA) to do the prediction. FWL and TWA have really closed relationship with hole number.

- $FWL:\ \ 0.0032 * hole\#^2 + 1.5575 * hole\# - 3.4767$
- $TWA:\ \ -0.1597 * hole\#^2 + 50.903 * hole\# - 27.345$

C. Dummy Module

Dummy module is using average value as the result of each prediction. We use dummy modules to tell if a model works, After we plot the violin plots, compared with other methods, the 20 components PCA on (thrust force and torque) has the best performance.

V. CONCLUSIONS

From Fig 14, compared with other methods, LR at 20-PCA and force + torque FFT are the most effective for predicting roughness coefficients. Secondly, the prediction of LR in the hole number has exceeded 80%. Combining the hole number provided by Professor Dave with the tool ware, we can also accurately predict the type of tool ware. Till now, most of our experiments are single-layer neural networks. Multi-layer neural networks have better results, such as LeNet-5 in CNN. This indicates that we should try to predict the roughness coefficient which can show the quality of drill holes using deep learning in the next stage.

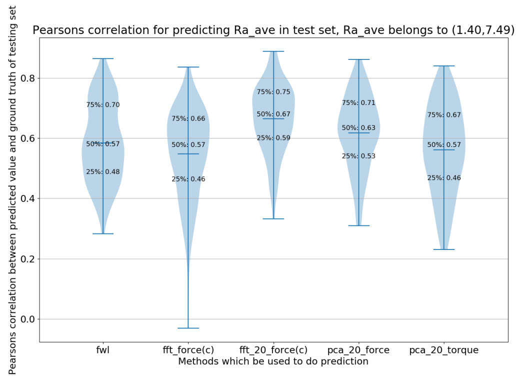


Fig. 13: 200 times violin plot on Tool ware, PCA and FFT

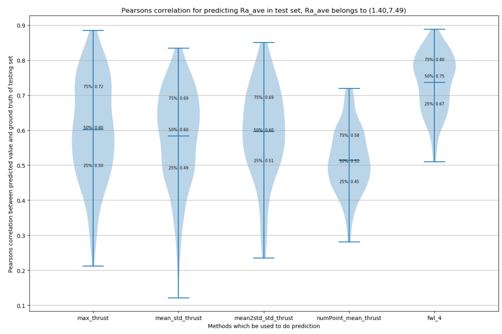


Fig. 14: 200 times violin plot on force character combinations

REFERENCES

[1] Bhatt, P. and Goe, A. (2017). Carbon Fibres: Production, Properties and Potential Use. Material Science Research India, 14(1), pp.52-57.

[2] J. Cheng, J. Qiu, X. Xu, H. Ji, T. Takagi, and T. Uchimoto, "Research advances in eddy current testing for maintenance of carbon fiber reinforced plastic composites," International Journal of Applied Electromagnetics and Mechanics, vol. 51, no. 3, pp. 261–284, Aug. 2016.

[3] Haytham Fayek, "Speech Processing for Machine Learning: Filter banks, Mel-Frequency Cepstral Coefficients (MFCCs) and What's In-Between," Haytham Fayek, 21-Apr-2016. [Online]. Available: https://haythamfayek.com/2016/04/21/speech-processing-for-machine-learning.html. [Accessed: 12-Dec-2019].

[4] Y. Lecun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," Proceedings of the IEEE, vol. 86, no. 11, pp. 2278–2324, 1998.

[5] M. Bleyberg, "A categorical entity-relationship model of databases," [Proceedings] 1991 Symposium on Applied Computing.

[6] M. Visvalingam and J. D. Whyatt, "The Douglas-Peucker Algorithm for Line Simplification: Re-evaluation through Visualization," Computer Graphics Forum, vol. 9, no. 3, pp. 213–225, 1990.

[7] K. Ito and K. Xiong, “Gaussian Filters for Nonlinear Filtering Problems,” 1999.
[8] X. Qiu, P. Li, Q. Niu, A. Chen, P. Ouyang, C. Li, and T. J. Ko, “Influence of machining parameters and tool structure on cutting force and hole wall damage in drilling CFRP with stepped drills,” The International Journal of Advanced Manufacturing Technology, vol. 97, no. 1-4, pp. 857–865, 2018.