CNC Milling Machine Anomaly Classification with Continual Active Learning

Eden Kim, Seungchul Son, Seokkap Ko
Honam Research Center (HRC)
Electronics and Telecommuncations Research Institute (ETRI)
Gwangju, Korea
{kimed93, sson, softgear}@etri.re.kr

***Abstract*—Computer Numeric Control (CNC) milling machines on the manufacturing field has significantly impact on product quality. Anomalies during CNC milling can disrupt manufacturing costs. Addressing such failures manually presents substantial challenges. As an alternative, artificial intelligence (AI) has been used manufacturing sites, enabling real-time data analysis and anomaly detection. However, AI's realization relies on quality data acquisition, a time-intensive process compounded by labeling. We proposes a continual active learning framework to enhance AI model learning from limited datasets in manufacturing. This paper presents three active learning methodologies: Least Confidence (LC), Entropy Sampling (ES), and Active Transfer Learning for Adaptive Sampling (ATLAS). We evaluate these methods using CNC milling machine data against Simple Random Sampling (SRS). Results highlight those methods have improved performance rather than SRS.**

***Index Terms*—CNC milling machine, Manufacturing, Anomaly Classification, Deep Learning, Continual Active Learning**

I. Introduction

The Computer Numeric Control (CNC) milling machine situated within manufacturing sites plays an important role and has a great influence on the quality of the product [1]. The occurrence of anomaly situations within the CNC milling process can significantly impact cost problems of the manufacturing site. However, the consistent vigilance and management of such failure scenarios by human personnel present formidable challenges. Consequently, the concept of smart factory has emerged as an imperative paradigm within the contemporary manufacturing landscape [2]. This trajectory is intrinsically aligns with the forefront of the Fourth Industrial Revolution, wherein smart factories emerge. These factories hold the potential to not only elevate productivity levels but also enhance the overall quality of produced items. Of particular significance is the convergence of Artificial Intelligence (AI) advancements, which have imbued the smart factory concept with innovate capabilities. The integration of AI and machine learning technologies has enabled real-time data analysis and pattern recognition within production processes [3]. It is important to underscore, however, that for the realization of an AI-powered automated system as outlined, the acquisition of ample high quality of data is indispensable [4]. The process of data accumulation for practical application is both time-intensive and resource-demanding included further by the laborious task of data labeling. This intricate interplay underscores the necessity of pragmatically implementing AI within the manufacturing sphere, necessitating a continuous cycle of data collection, labeling, and model refinement.

In this paper, we propose an continual active learning framework that can continuously learn AI models like online-model in the manufacturing field from small dataset. Active learning is the approach that involves an interactive process where the model actively selects and queries the most informative data points from a larger pool of unlabeled data [5]. Three selected methodologies of active learning: Least Confidence (LC), Entropy Sampling (ES) and Active Transfer Learning for Adaptive Sampling (ATLAS) was described, and verified with CNC milling machine dataset compared with Simple Random Sampling (SRS) methods. Overall, LC, ES, and ATLAS methods all showed better performance than SRS. Among them, ATLAS in particular showed the best performance.

The remaining paper is organized as follows: In section II, we describe the proposed methodology about anomaly classification with continual active learning. Section III shows the experimental results of proposed methods. Finally, conclusion will be shown in section IV.

II. Methodology

The primary goal of continual active learning is to make the learning process more efficient by focusing on the most relevant examples that are likely to provide the most information to the model. Fig. 1 shows the process flow of continual active learning. It assumes that data collection is continuously taking place. The collected data belongs to an unlabeled dataset, and an active learning module is formed based on the classification model. The queries x^* selected by the active learning module are transmitted to oracle for labeling, and the configured $\{x^*, y\}$ data is included as a trainable dataset, and used for retraining the classification model.

A. Classification model

A fully connected neural network was selected as a model for anomaly classification from the collected manufacturing data. The model for classification anomality can be replaced with RNN-based models or CNN-based models to improve better performance depending on the situation. In this paper, since we want to confirm that the accuracy of the model is

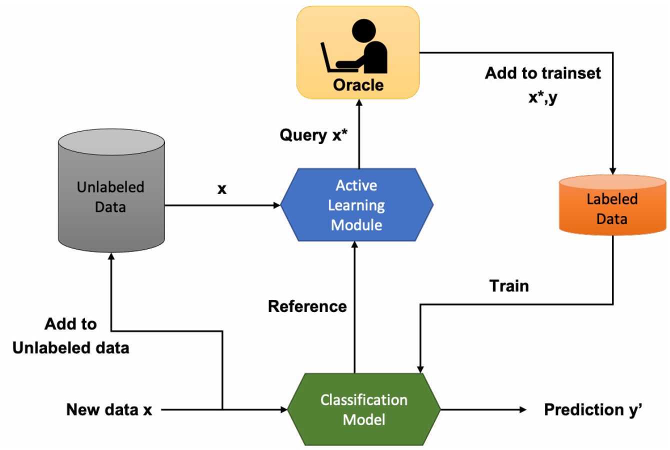

Fig. 1. Continual Active Learning Framework

```
Layer (type)                 Output Shape              Param #
=================================================================
flatten_1 (Flatten)          (None, 3500)              0

dense_3 (Dense)              (None, 128)               448128

dense_4 (Dense)              (None, 64)                8256

dense_5 (Dense)              (None, 2)                 130

=================================================================
Total params: 456,514
Trainable params: 456,514
Non-trainable params: 0
```

Fig. 2. Anomaly Classification Model Structure

improved through continuous retraining starting from a limited amount of data, we adopted a basic deep learning method rather than other complex models. This model is consisted of four layers as shown in Fig. 2, using ReLU activation functions, and softmax function at the end.

B. Active Learning module

In this section, three methods are introduced as active learning modules: LC, ES, and ATLAS. Moreover, SRS was used as sampling without replacement, and was compared with other active learning methods as the most basic method for data sampling.

Uncertainty Sampling - Least Confidence (LC). For choosing the most informative samples, uncertainty sampling identify unlabeled data that are the near a decision boundary in classification model [6]. In case of least confidence method, the data with the lowest maximum probability should be selected. Eq. (2) shows the maximum probability $\hat{y}$ among each class confidence. Finally, sample s_{LC} is chosen by the lowest value within maximum probability as shown Eq. (1). The advantage of LC is that it is easy to implement, but it is vulnerable to outliers.

$$s_{LC} = \underset{x}{argmax}(1 - P(\hat{y}|x)) \tag{1}$$

$$\hat{y} = \underset{y}{argmax}(P(y|x)) \tag{2}$$

Uncertainty Sampling - Entropy Sampling (ES). As different as LC method, entropy sampling use the maximum entropy as a uncertainty measure indicator instead of confidence [6]. Eq. (3) shows how to get sample s_E with maximum entropy. Maximum entropy is an indicator of the amount of information and often shows the strongest performance among uncertainty sampling. On the other hand, it is also vulnerable to outliers like LC method.

$$s_E = \underset{x}{argmax}(-\sum_i P(\hat{y}_i|x) \log P(\hat{y}_i|x)) \tag{3}$$

Active Transfer Learning for Adaptive Sampling (ATLAS). Unlike the LC and ES methods mentioned above, which measure uncertainty from confidence results of classification model, active transfer learning selects the samples with learning a new model by shifting the existed classification model [7]. To learn the new model, validation label data is manipulated as follows. If the result of classification is correct, validation label data become 0 (correct), and if the result is inaccurate, validation label data become 1 (incorrect). The ATLAS model is configured by changing the layer at the end through transfer learning based on the existing trained classification model, and trained with manipulated validation data. The confidence value with the closest incorrect is determined to be the most informatively necessary and is selected as the sample $\{s_{ATL}\}$ that needs to be newly labeled. ATLAS has the strength of being flexible through continuous updates without being vulnerable to outliers. However, there is a disadvantage that it takes a long time for continuous learning.

III. Experimental Study

A. Experimental Settings & Dataset

To verify feasibility of active learning methods, we construct the experiments with SRS sampling and three active learning methods. The CNC milling machine dataset [8] used in this experiment was collected in the field. Totally, seven variables are collected: CNC cutting conditions-S, CNC cutting conditions-F, X tool position, Y tool position, Z tool position, spindle motor-U CT and spindle motor-V CT. The dataset was preprocessed with 500 time steps for anomaly classification.

Pre-processed datasets are divided into training set $\{X_{tr}, Y_{tr}\}$, validation set $\{X_{val}, Y_{val}\}$, test set $\{X_{te}, Y_{te}\}$ and unlabeled set $\{X_{un}, Y_{un}\}$. The ratio of sets is decided 1:3:3:3 because it assumes that training starts from small amount of data. The experiment scenario is as follows.

1) Classification model is trained with labeled training set $\{X_{tr}, Y_{tr}\}$.
2) The accuracy of classification model is calculated with test set $\{X_{te}, Y_{te}\}$.
3) The methods of the active learning module (SRS, LC, ES, ATLAS) are based on the classification model results on the validation set $\{X_{val}, Y_{val}\}$.
4) Top number of N among unlabeled data $\{X_{un}, Y_{un}\}$ in the active learning module are selected new samples for labeling.

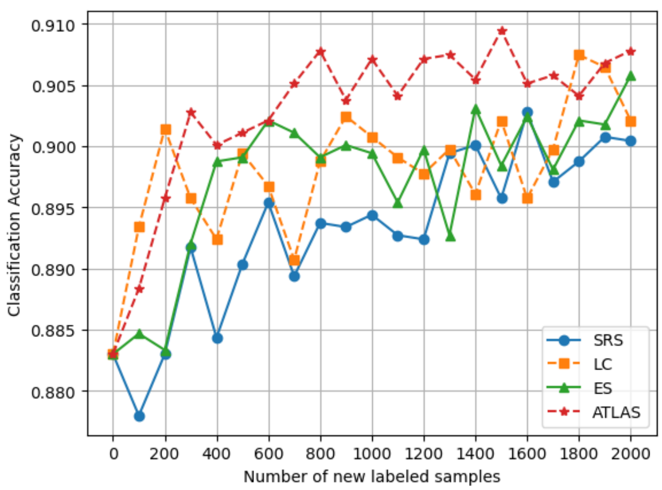

Fig. 3. Experimental Results

5) Selected new sample is labeled and updated the training set $\{X^*_{tr}, Y^*_{tr}\}$.
6) Classification model is retrained with updated labeled training set $\{X^*_{tr}, Y^*_{tr}\}$.
7) Go to step 2) and continue repeating for the set number of cycles.

In this experiments, the number of data is 994 in training set, 2983 in validation set, 2983 in test set, and 2983 in unlabeled set. The number of cycles is set 20, and the number of selected new samples in active learning module per cycle is set 100 based on a tenth of the training set.

B. Results

The experimental results are shown in Fig. 3. All of active learning module methods perform better than SRS method. Since SRS select samples randomly, the probability of adding example which is helpful to retrain is the lowest. On the other hand, LC method shows the fastest accuracy improvement in the beginning ($N \leq 200$). Though it performed lower classification accuracy rather than the other in a few section, it shows significantly good results 0.9075 at 18 epoch ($N = 1800$). ES method is gradually improved, and then it shows similar performance with LC, SRS in the second half ($1000 \leq N \leq 2000$). In 10-18 epochs ($1000 \leq N \leq 1800$), LC, ES, and SRS results is fluctuated because classification model is not enough to robust. The result of ATLAS shows robust and high accuracy. As a result, suggested active learning methods are verified to perform better than SRS for adapting in the field. Particularly, ATLAS method can be best solution since it is accurate and robust.

IV. Conclusion

In this paper, we proposed anomaly classification with continual active learning in manufacturing sites. A fully connected neural network is used for anomaly classification model, LC, ES and ATLAS methods are suggested for active learning. All of active learning modules show the better performance than SRS, which is applicable to real field. ATLAS particularly shows the fastest growth rate of accuracy. As a result, it is expected to be of great help in terms of operation and management by presenting a method to gradually solve the lack of data at manufacturing sites using CNC milling machines. As future work, we will conduct several experiments by changing the composition of the dataset in order to reduce fluctuation of result. We also plan to apply other unlabeled CNC dataset from other sites.

Acknowledgment

This research was supported by Energy AI Convergence Research & Development Program through the National IT Industry Promotion Agency of Korea(NIPA) funded by the Ministry of Science and ICT(No. 1711171060)

References

[1] R. Lins, B. Guerreiro, R. Schmitt, J. Sun, M. Corazzim, and F. Silva, “A novel methodology for retrofitting CNC machines based on the context of industry 4.0,” 2017 IEEE international systems engineering symposium (ISSE). pp. 1-6, 2017.

[2] B. Chen, J. Wan, L. Shu, P. Li, M. Mukherjee, and B, Yin, “Smart factory of industry 4.0: Key technologies, application case, and challenges,” Ieee Access, vol. 6, pp. 6505-6519, 2017.

[3] T. Wuest, D. Weimer, C. Irgens, and K. Thoben, “Machine learning in manufacturing: advantages, challenges, and applications,” Production & Manufacturing Research, vol. 4, no. 1, pp. 23-45, 2016.

[4] M. Tnani, M. Feil, and K. Diepold, “Smart data collection system for brownfield CNC milling machines: A new benchmark dataset for data-driven machine monitoring,” Procedia CIRP, vol. 107, pp. 131-136, 2022.

[5] B. Settles, “Active learning literature survey,” 2019.

[6] V. Nguyen, M. Shaker, and E. Hüllermeier, “How to measure uncertainty in uncertainty sampling for active learning,” Machine Learning, vol. 111, no. 1, pp. 89-122, 2022.

[7] X. wang, T. Huang, and J. Schneider, “Active transfer learning under model shift,” International Conference on Machine Learning, pp. 1305-1313, 2014.

[8] https://github.com/etri/Smart_Factory_CNC_Dataset