

Sliding Window Filter Based Strip Breakage Modelling for Failure Prediction

Jialu Yang, Zheyuan Chen, Ying Liu *Member, IEEE* and Michael Ryan

Abstract— In the production of cold-rolled strip products, strip breakage is one of the most common failures during the cold rolling process. However, the existing prediction models on strip breakage use the conventional sliding window algorithm to process the time series data collected from the actual production, resulting in a massive amount of non-informative data, which increases the computational cost for data-driven modelling. In order to tackle this issue, this article proposed a sliding window filter method to optimise the data pre-processing of the strip breakage. Firstly, based on the existing research and understanding of strip breakage, the data characteristics in the process of strip breakage was analysed. Based on the analysis, sample variance (VAR) and length normalised complexity estimate (LNCE) were chosen to determine how informative the time window was related to strip breakage. Secondly, compared with the conventional sliding window approach, the sliding windows were classified through a filter using VAR and LNCE. Thirdly, the filtered data was fed into the Recurrent Neural Network (RNN) for strip breakage modelling. An experimental study based on actual production data collected by a cold-rolled strip manufacturer was conducted to verify this method's effectiveness. The results show that pre-processing data using the sliding window filter decreases the model's computational cost.

I. INTRODUCTION

Cold rolling is an indispensable process in steel production. In this process, the steel strip enters the cold rolling mill, and its thickness is gradually reduced to the required size under the vast pressure of the rollers. In the production process of high silicon steel, cold rolling reduces the steel strip's thickness by 90% [1]. Because pressure on the steel strip and the amount of deformation during this process are enormous, strip breakage often occurs. Strip breakage (strip snap) is one of the most frequent failures in cold rolling [2]. This failure not only damages the working equipment and steel strip but also causes long-term work stagnation and ultimately reduces production efficiency [3].

There have been many studies investigating strip breakage causes in recent years. However, these mainly focus on analysing the problem using metallurgy or mechanical theory [2, 4]. With the widespread use of sensors in the production process, it is now possible to use data analysis to detect and prevent failures [5]. Meanwhile, with the continuous development of information technology, analysis of big data has become very common [3, 6]. In large-scale industrial production such as the steel industry, data analysis has begun

to be widely used. However, because the amount of data collected in production is tremendous and contains a massive amount of irrelevant data, data analysis often requires a lot of computing time. At the same time, the data collected by sensors is usually time-series data. For the analysis of this type of data, studies using the sliding window algorithm were proposed. However, to better reflect the data's characteristics, these studies have selected many variables, which further increase the complexity of the data, thereby increasing the computing time and complexity.

With the aim of analysing big data in industrial production and eliminate irrelevant data, this paper proposes a method of pre-processing data using a sliding window filter to reduce the computing time. Regarding the failure analysis, typically, the amount of failure data is often much less than that of normal operation data, which results in a data imbalance problem. In previous studies, the sliding window algorithm was used to solve the data imbalance problem [7]. By adding a sliding window filter, the balance of data can be better improved [8]. Since the failure occurs suddenly, it will bring a sudden change in the data. Therefore, selecting a suitable filter based on this mutation can effectively eliminate data with nothing to do with the fault and reduce the total data amount [9].

The structure of this article is as follows. In Section 2, a review of strip breakage in cold rolling and strip breakage modelling is addressed, followed by a review on using the sliding window for data filtering. Subsequently, the methodology of using a sliding window filter to modelling and analysing strip breakage in the cold rolling process is proposed in Section 3. A case study using real-world cold rolling process data is reported in Section 4. This case study's primary purpose is to evaluate through experiments whether the performance of data analysis and computing time has been improved after using the sliding window filter. Finally, Section 6 presents the conclusions and future work of this article.

II. LITERATURE REVIEW

A. Strip breakage in cold rolling

Cold rolling is a steel manufacturing process. This process's primary purpose is to reduce the thickness of the steel strip entering the machine to achieve a required thickness [4]. This process can create anisotropy for the steel in the rolling direction and increase the material's strength by 20% [10].

Moreover, this method can improve the material's surface roughness, thereby increasing the material's dimensional tolerance. Compared with hot rolling, cold rolling can produce thinner and more precise plates, bars and materials.

Cold rolling processes the incoming strip into a thinner material by applying tremendous pressure. However, due to the enormous stress imposed on the material, sometimes there is an uneven force problem. Simultaneously, the material undergoes massive deformation during this process, so the strip often breaks. This phenomenon is called strip breakage [11].

This failure will have a severe impact on the production process. First of all, due to the steel's fracture, the steel's mechanical properties are significantly reduced, which cannot meet the production requirements, and the entire steel coil needs to be reheated, which causes serious waste [12]. Secondly, when a strip breakage occurs, the steel plate subjected to enormous pressure and tension will instantly fly out of the rolling mill, which will cause serious safety hazards and cause severe damage to work equipment. To ensure smooth production, after the strip breakage occurs, the factory needs to conduct a comprehensive inspection of all equipment to remove hidden dangers, which further reduces production efficiency [13].

Since strip breakage is the most common failure in the cold rolling process, there have been many studies on its causes. These studies mainly divided the causes of strip breakage into four categories: equipment failure, improper setting of rolling parameters, material defects and improper operation [14]. These studies identified equipment failure as the leading cause of strip breakage. In the production process, when the servo valve fails, unbalanced tension will be generated between the machine frames, causing the steel strip to be squeezed and broken [15]. Rolling parameters will have a significant impact on the entire process. For example, the feed parameters in the cold rolling process will seriously affect the smoothness of the material, which is also one reason for strip breakage. Material defects are mainly reflected in flatness defects. When the steel enters the rolling equipment, if its flatness is not high enough and the material cannot wholly contact the rolling mill, shifting, strip breakage, and axial forces will occur. At present, most of the rolling mills used in production are automated. In this case, there are fewer strip breakages caused by improper operation. However, some failures caused by workers not aligning the steel strip into the machine have also occurred.

B. Previous studies on strip breakage modelling

In the past few decades, many models of cold rolling processes have been developed and proposed. Most of the earliest cold rolling process models are constructed based on rolling parameters, such as rolling force, strip tension and some operating parameters [16]. The cold-rolling model based on rolling parameters developed by Orowan is a representative. These models analyse the productivity, quality and failures in cold rolling based on the conventional rolling force formula [10].

However, due to the cold rolling process's complexity, especially with the widespread application of sensors, more and more relevant data collected during the cold rolling process, the reasonable approximation provided by the

conventional rolling force formula can no longer meet the requirements [17]. Some researchers have begun to apply data analysis methods to strip breakage modelling in recent years [18]. The data collected in production, such as friction conditions, roll pressure, and temperature, often vary greatly, so theoretical analysis is challenging, and it is complicated to analyse these factors using formulas [19].

In some models, the researchers integrated rolling principle analysis and data analysis methods. First, in data pre-processing, the relevant features are selected according to the principle. Then further data analysis is used to obtain the most relevant features of strip breakage, and finally, a model is established based on these features [20]. This method was used to determine the 20 features that are most relevant to strip breakage during cold rolling [21].

The author of this paper also used this method to establish a strip breakage analysis and prediction model in the previous study [19]. The model is based on the analysis of the existing features. According to the various data analysis methods used, different features are selected for modelling, and the results are compared.

C. Using the sliding window for time-series data analysis

In the failure detection research, because the amount of failure data is less than that of normal data, a skewed distribution exists. A skewed distribution means that there is much more data in a specific class in the data set than in other classes [22]. This distribution can lead to severe data imbalance problems [23]. Simultaneously, since data collected in actual production are usually sampling points, not time series data, it will bring a lot of inconvenience in the analysis process [24]. Therefore, to reduce the amount of data, solve the data imbalance problem, and convert the data into time-series data, the sliding window algorithm is used by some researchers.

The method of combining several specific data in the raw data set into a new window to reduce the total amount of data is called the window algorithm [25]. When studying multivariate time series data, researchers usually integrate several continuous data into a new set of windows to better study the data's characteristics. These continuous data's characteristic values, such as maximum value, minimum value and sum, will become the new window's value [26].

The sliding window algorithm is a method of processing and transforming data one by one through a certain window

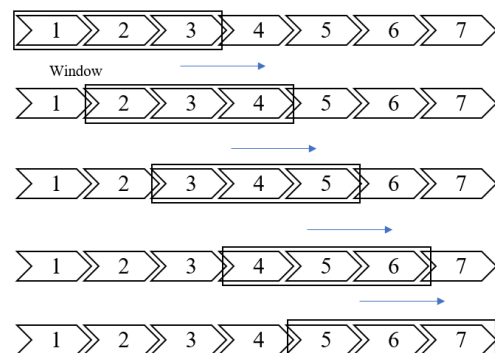


Figure 1. Sliding Window Technique

size. As illustrated in Fig.1, assuming that the given window size is 3, the window starts from the first element of the data group and slides to the right one element at a time. It can be found that each element in the raw data will appear three times (the window length), so it will reflect the raw data's characteristics well [27].

III. METHODOLOGY

With the aim of reducing the total amount of data and computing time, a sliding window filter is used to filter the data further, and model analysis is established based on the processed data. First of all, this article selects the appropriate filter based on the actual situation of strip breakage. Then, the time series data is classified using a sliding window filter. Finally, the RNN-based data analysis model is established and evaluated. The flow chart of proposed method is shown in Figure 2.

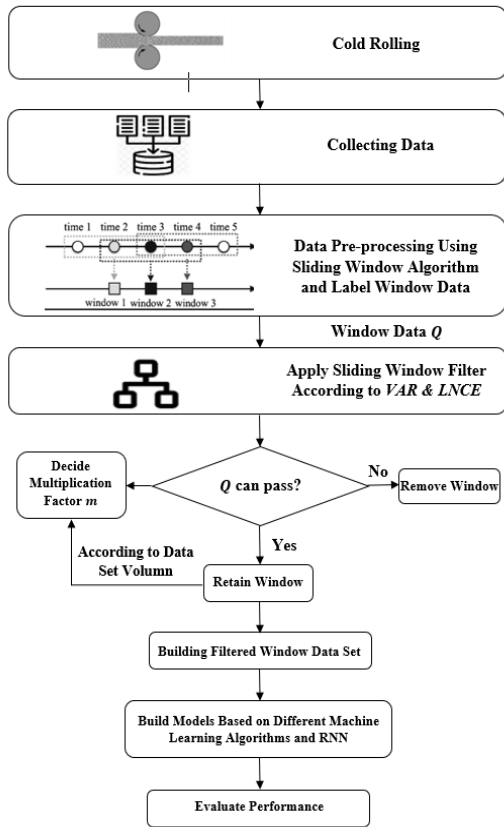


Figure 2. Methodology flowchart

A. Construction of the sliding window filter

Usually, the conventional sliding window needs to consider all the measured values in the data set. The number of windows will change with the window size and window step length, which will significantly increase the calculation amount.

Fig.3 is part of the data collected during the cold rolling process of a factory. The features shown in the figure are the most relevant features of strip breakage based on previous studies [28]. Among them, when Fast Stop changes from 0 to

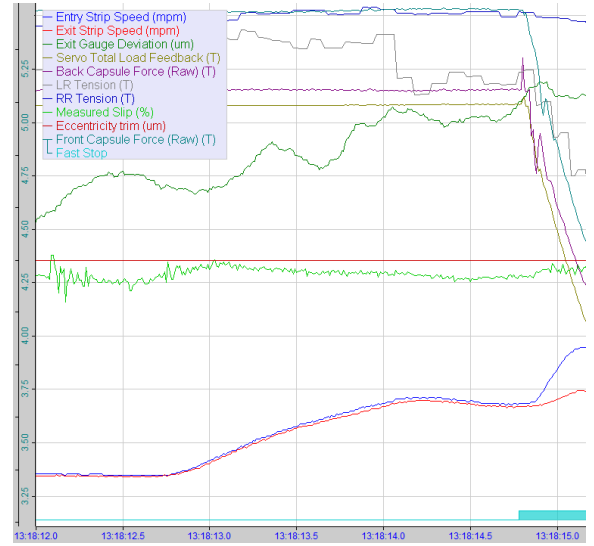


Figure 3. Example of cold rolling process data before strip breakage (ibaAnalyzer)

1, that is, when the red line at the bottom has a sudden change, it represents strip breakage. It can be found that before the strip breakage occurs, these features have begun to fluctuate sharply. According to the failure principle analysis, the occurrence of strip breakage is damage to the steady-state, so there must be parameter fluctuations. Therefore, the data fluctuation size can be used as an index to filter out the data most relevant to strip breakage. Similarly, this method can also be used in other failure predictions.

To reduce computing time, solve the data imbalance problem, and comprehensively analyse the failure data in the manufacturing process, this paper proposes a sliding window filter, which can effectively remove data that is not related to failure from the raw data and only retains failure related data. From this point of view, this filter can be regarded as a binary classifier. This classifier divides the data into two categories, related and irrelevant, according to selected features. Many filter candidates, including statistical measures, such as the sample variance (VAR) and length normalised complexity estimate (LNCE), are suitable to achieve this goal [29].

When given a time series $Q = (q_1, \dots, q_n)$ with length n , VAR and LNCE can be defined as follows:

$$VAR(Q) = \frac{1}{n} \sum_{i=1}^n (q_i - \mu)^2 \quad (1)$$

$$LNCE(Q) = \frac{1}{n-1} \sqrt{\sum_{i=1}^{n-1} (q_i - q_{i+1})^2} \quad (2)$$

Among them, μ is the average value of this time series data. VAR and LNCE can respectively represent the degree of change of the window itself and the difference between it and the overall trend. After defining the statistical indicators, the statistical results of VAR and LNCE for all training sets can be obtained to determine the appropriate range for classification.

Since all windows will be calculated and classified by this classifier, to avoid the possible excessive pruning of relevant data, a multiplication factor m will be introduced in Chapter IV, which allows the filter to expand the filtering range.

B. Classifying time series data using sliding window filter

After determining the filter index, the time series data were screened, first, converting the data collected in production into windows and then filtering each set of data and preliminarily classifying them into relevant and irrelevant. After that, by comparing the size of the two sets, the appropriate multiplication factor m was determined.

C. Modelling strip breakage based on RNN

This study uses the Recurrent Neural Network (RNN) algorithm to establish a data analysis model to verify the proposed methodology's effectiveness. RNN is a recurrent neural network that takes sequence data as input, recursively in the direction of sequence evolution, and all nodes (cyclic units) are connected in a chain [30]. Because RNN can retain the recent memory of input data and fully mine the information of multivariate time series data, it is widely used in time series processing [21].

IV. CASE STUDY

The data for this case study was provided by a high-silicon electrical steel production factory. In the plant's cold rolling process, the steel strip is rolled back and forth five times through a reversible rolling mill. Since this process will reduce the thickness of high silicon steel by 90%, strip breakage often occurs. The factory engineer marked the coil data where the strip breakage occurred. The author of this conference paper has used some data analysis methods to analyse the data and established a model that can be used to predict strip breakage in the previous study [21]. However, since the methods used are relatively rudimentary, and the sliding window does not entirely solve the data imbalance problem, these experiments are only for reference.

The purpose of establishing this experiment is to prove that the sliding window filter can reduce the computing time of

model building and improve the data imbalance problem. With the aim of quantifying the decrease in computing time, this study compared the time it takes to train a model using the RNN-based algorithm with and without the filter. Also, this article introduces the area under the ROC curve (AUC) to evaluate the improvement effect of the filter on the data imbalance problem. This indicator is widely used to compare and evaluate the data imbalance of models [31].

As mentioned above, Fast Stop can be used to determine whether a strip breakage has occurred. However, since this article aims to establish a failure prediction model, the data that the established model needs to predict should be before the failure occurs. To achieve this, the sampling points within 0.5 seconds before the strip breakage occurs is marked as the Pre-Snap point, and a model for analysing the Pre-Snap data is constructed.

A. Data collection and description

The data used in this case study is the actual production data collected by Production Data Acquisition (PDA) equipment. PDA is a collection system that can collect more than 1,000 cold rolling parameters, including setting values, operating variables and variables in the production process. These data are collected 100 times (100 sampling points) per second in real-time.

Although the collected data contains more than 1,000 process variables in the cold rolling process, according to the previous study [21], 11 variables that are most relevant to cold rolling have been selected, as shown in Table 1.

B. Data pre-processing

The collected data is the raw time-series data from the PDA system directly. These raw data need to transfer into windows before applying the sliding window filter. In each coil of steel strip, all the data in the 5 seconds before the stirp breakage happens is collected. As mentioned earlier, each set of data contains 500 sampling points, and 50 of them are marked as Pre-Snap points. This experiment collected 137 coils during the year 2017 and 2018. Among these coils, 100 were used as the training set, and the remaining were used as the test set.

First, the sliding window algorithm is applied to the raw data. This article compares the windows obtained by setting the window length from 5 to 500. After comparing different window length settings, the imbalance problem of the new window data, and the degree of response to the raw data characteristics, 200 was chosen as the window length. After applying the sliding window algorithm, 50 sets of windows

TABLE I. SELECTED VARIABLES

Variable Name	Description
Entry Strip Speed(m/min)	Speed of steel strip entry
Exit Strip Speed(m/min)	Speed of steel strip left
Exit Gauge Deviation(mm)	At the exit of the machine, the thickness deviation of the steel coil
Eccentricity trim(mm)	Correction parameters used to keep the roller circular
Servo Total Load Feedback	Pressure on steel coil
Front Capsule Force (Raw)	Force on the strip coil of the front capsule
Back Capsule Force (Raw)	Force on the strip coil of the back capsule
LR Tension	The tension of the strip on the left side of the machine
RR Tension	The tension of the strip on the right side of the machine
Fast Stop	Machine emergency stop status indication 1: Emergency stop 0: Normal operation
Measured Slip (%)	Displacement ratio between strip coil and machine

TABLE II. WINDOW FEATURES

Sum	The sum of each feature in the window
Mean	The average of each feature in the window
Min	The minimum of each feature in the window
Max	The maximum of each feature in the window
Median	When the feature values are arranged in order of size, the middle value
Std	The standard deviation of each feature

contain Pre-Snap points, marked as Pre-Snap data, and 251 windows do not contain this point and are marked as Normal data. Also, to use the conventional machine-learning algorithm to build models, this article extracts window features. This paper extracts six window features that best represent raw data characteristic, as shown in Table 2.

Two different window sets are established, filtered and unfiltered, to study the sliding window filter's effect on the model.

The first set of windows will pass through the sliding window filter, and the VAR and $LNCE$ of each roll of steel strip data (marked as VAR_1 and $LNCE_1$) and in each window (marked as VAR_2 and $LNCE_2$) are calculated, respectively. Since VAR and $LNCE$ can respectively reflect the fluctuation difference between windows and raw data, it is necessary to synthesise the two indices for judgment.

If the VAR and $LNCE$ of the window are more significant than the raw data set, these window data are considered to be related to strip breakage. Otherwise, they are irrelevant to strip breakage.

However, because the windows that can pass through this filter are too few in the application process, there is an excessive pruning problem. Therefore, this article introduces multiplication factor m to increase the amount of data passing through the filter. Since this experiment selected 11 various features, it was found that when multiple features are filtered, the amount of data passing through the filter will be reduced. This article compares the total amount of data obtained when different numbers of features pass the filter and finally determines that when any three or more features can pass the filter, the data is retained. After comparing the filtering performance of the filter under different parameters, 0.8 is selected as the value of m . When m is less than 0.8, the filter cannot filter irrelevant data. When m is greater than 0.8, the filter will also remove part of the Pre-Snap data.

Finally, the filtering rules used to determine the filter group data are as follows:

$$VAR_2 > m \times VAR_1 \quad (3)$$

$$LNCE_2 > m \times LNCE_1 \quad (4)$$

The ratio of Pre-Snap windows to Normal windows is 50:251 in the raw data. After applying the filter, this ratio is

reduced to 6802:17927. At the same time, the overall data volume has been reduced by nearly 40%. The second set of windows did not pass the filter and was directly used for modelling.

C. Experiments

In order to study the influence of the filter on the model, two experimental scenarios are designed in this Case Study. The first experimental scenario is based on conventional machine learning algorithms, using unfiltered and filtered data to build predictive models and compare various indicators of these models. The second scenario of experiments use RNN-based algorithms to build models, and the performances are evaluated.

In the first scenario, this article uses three conventional base classifiers to build models, which are J48, IBk and NaiveBayes. Besides, four ensemble classifiers, Random Forest, Bagging, AdaBoost and Vote, are also used to build models. Random Forest is the sub-model algorithm of Adaboost and Vote, which is the best algorithm selected based on previous experiments.

In the second scenario, three RNN-based algorithms, RNN, LSTM, and GRU, are used to build predictive models. The advantage of RNN is that it can directly analyse the raw time-series data. However, since the current experiment data comes from multiple rolls of different steel strips, it is necessary to perform sequence to vector conversion on these data before they can be put into RNN for modelling. Each window data is regarded as a vector, and the overall data is analysed. The RNN model's EPOCH is set to 200, and the BATCH SIZE is set to 128 to better compare the improvement of computing time after applying the filter.

D. Results and discussions

The results of various data analysis models are shown in Table 3. It can be found that when analysing the windows after applying the filter, for most conventional machine learning algorithms, the data imbalance problem has been improved to a certain extent. The AUC of these models has been improved to varying degrees. At the same time, it can also be found that the time spent on training the model has been significantly reduced. This improvement is apparent in the construction of the RNN-based models.

TABLE III. RESULTS OF THE EXPERIMENTS

Data Entry	Training scheme	Algorithm	ACC	AUC	Computing Time	Data Entry	Training scheme	Algorithm	ACC	AUC	Computing Time
Raw Sliding Windows	Extracted features + conventional base classifier	J48	0.796	0.704	2.08s	Filtered Windows	Extracted features + conventional Base classifier	J48	0.767	0.621	1.55s
		IBk	0.815	0.619	63.84s			IBk	0.718	0.647	21.98s
		NaiveBayes	0.759	0.719	0.26s			NaiveBayes	0.748	0.739	0.12s
	Extracted features + ensemble learning	Random Forest	0.856	0.820	8.44s		Extracted features + ensemble learning	Random Forest	0.819	0.844	4.73s
		Bagging	0.821	0.790	6.49s			Bagging	0.754	0.782	3.37s
		AdaBoost	0.872	0.817	8.54s			AdaBoost	0.820	0.832	4.63s
		Vote	0.860	0.819	8.16s			Vote	0.812	0.840	4.47s
	Time window + RNN-Based model	RNN	0.886	0.813	7451.21s		Time window + RNN-Based model	RNN	0.873	0.853	3325.30s
		LSTM	0.912	0.819	7324.86s			LSTM	0.906	0.846	3168.13s
		GRU	0.906	0.835	7550.92s			GRU	0.886	0.886	3360.21s

However, it can be found from the results that the accuracy of these models has been partially reduced. There are two main reasons. Firstly, the filter reduces the amount of overall data, leading to a decrease in accuracy. The second is that the amount of data used in this experiment is not large enough, so this change in the amount of data has an obvious manifestation in the model's performance. This accuracy decrease is acceptable, considering the use of filters to improve the modelling rate.

Moreover, it should be noted that for the RNN-based model, the accuracy drop caused by the sliding window filter is minimal. Since conventional machine learning algorithms usually require a short computing time, the sliding window filter cannot achieve a good performance improvement effect for models based on these algorithms. This filter is more effective for algorithms similar to RNN, whose accuracy is less affected by the total amount of data and has a long computing time.

V. CONCLUSION

In this work, a sliding window filter is proposed, which can improve the construction of failure prediction models. This filter compares the fluctuations of the window data and the original data, filters out the window data that is most relevant to the fault, and eliminates irrelevant data, thereby reducing computing time and improving data imbalance.

The final test results show that this filter can effectively reduce the modelling process's computing time and improve most algorithms' data imbalance problem. For conventional machine learning algorithms, this method is not very practical. These algorithms have a relatively short operation time, and for some algorithms, the data imbalance problem after filtering is even more severe. For RNN-based algorithms, this filter has an excellent effect. It dramatically reduces the computing time and improves the data imbalance problem and does not cause a significant drop in accuracy.

In future work, comparison and evaluation of more filter methods are planned to find a sliding window filter that is more suitable for the failure prediction model.

REFERENCES

- [1] S. Iwadoh and T. Mori, "Effect of work roll materials and progress of manufacturing technology on cold rolling and future developments in Japan," *ISIJ international*, vol. 32, no. 11, pp. 1131-1140, 1992.
- [2] M. Mashayekhi, N. Torabian, and M. Poursina, "Continuum damage mechanics analysis of strip tearing in a tandem cold rolling process," *Simulation Modelling Practice and Theory*, vol. 19, no. 2, pp. 612-625, 2011.
- [3] K. M. Takami, J. Mahmoudi, E. Dahlquist, and M. Lindemo, "Multivariable data analysis of a cold rolling control system to minimise defects," *The International Journal of Advanced Manufacturing Technology*, vol. 54, no. 5-8, pp. 553-565, 2011.
- [4] X. Cui and A. Zhao, "Analysis the Causes of Deviation Breaking of Cold Rolling Strip," *Xinjiang Steel*, vol. 1, no. 2, pp. 44-47, 2013.
- [5] B. Liu, "Study and application of decoupling control of roll tilt and rolling force difference to inhibit strip break during cold rolling," *Metallurgical Industry Automation*, vol. 39, no. 6, pp. 40-44, 2015.
- [6] J. Xu, "Cause Analysis for Strip-breaking of Non-oriented Silicon Steel," *Jiangxi Metallurgy*, no. 05, pp. 26-28, 2015.
- [7] G. Lemaitre, F. Nogueira, and C. K. Aridas, "Imbalanced-learn: A python toolbox to tackle the curse of imbalanced datasets in machine learning," *The Journal of Machine Learning Research*, vol. 18, no. 1, pp. 559-563, 2017.
- [8] C.-H. Lee, C.-R. Lin, and M.-S. Chen, "Sliding-window filtering: an efficient algorithm for incremental mining," in *Proceedings of the tenth international conference on Information and knowledge management*, 2001, pp. 263-270.
- [9] G. Lesti and S. Spiegel, "A Sliding Window Filter for Time Series Streams," in *IOTSTREAMING@PKDD/ECML*, 2017.
- [10] E. Orowan, "The calculation of roll pressure in hot and cold flat rolling," *Proceedings of the Institution of Mechanical Engineers*, vol. 150, no. 1, pp. 140-167, 1943.
- [11] F.-x. Hou, J. Zhang, J.-g. Cao, and X.-l. Shi, "Review of chatter studies in cold rolling," *Gangtie Yanjiu Xuebao (Journal of Iron and Steel Research)*, vol. 19, no. 10, pp. 6-10, 2007.
- [12] K. Günther, G. Abbruzzese, S. Fortunati, and G. Ligi, "Recent Technology Developments in the Production of Grain - Oriented Electrical Steel," *steel research international*, vol. 76, no. 6, pp. 413-421, 2005.
- [13] K. Yan and H. Li, "Causes and Countermeasures of Cold Strip Break in Masteel," *Anhui Metallurgy*, vol. 1, no. 4, pp. 36-37, 2006.
- [14] S. Kasai, M. Namikawa, and T. Hiratani, "Recent progress of high silicon electrical steel in JFE steel," *JFE Steel Corporation, Tokyo, Japan, JFE Tech. Rep.*, vol. 21, 2016.
- [15] C. Andrei, M. Bogdan, N. Georgian, P. Ionut, and B. Zorica, "Cold rolling shape defects of stainless steel wide strips," in *European Conference of Chemical Engineering, ECCE*, 2010, vol. 10, pp. 154-158.
- [16] M. Kano and Y. Nakagawa, "Data-based process monitoring, process control, and quality improvement: Recent developments and applications in steel industry," *Computers & Chemical Engineering*, vol. 32, no. 1-2, pp. 12-24, 2008.
- [17] Q. Liu, T. Chai, H. Wang, and S.-Z. J. Qin, "Data-based hybrid tension estimation and fault diagnosis of cold rolling continuous annealing processes," *IEEE transactions on neural networks*, vol. 22, no. 12, pp. 2284-2295, 2011.
- [18] K. Kira and L. A. Rendell, "The feature selection problem: Traditional methods and a new algorithm," in *Aaai*, 1992, vol. 2, pp. 129-134.
- [19] Z. Chen, Y. Liu, A. Valera-Medina, and F. Robinson, "Strip Snap Analytics in Cold Rolling Process Using Machine Learning," in *2019 IEEE 15th International Conference on Automation Science and Engineering (CASE)*, 2019, pp. 368-373: IEEE.
- [20] M. Christ, A. W. Kempa-Liehr, and M. Feindt, "Distributed and parallel time series feature extraction for industrial big data applications," *arXiv preprint arXiv:1610.07717*, 2016.
- [21] Z. Chen, Y. Liu, A. Valera-Medina, and F. Robinson, "Characterising strip snap in cold rolling process using advanced data analytics," *Procedia CIRP*, vol. 81, pp. 453-458, 2019.
- [22] J. Brownlee, "A Gentle Introduction to Imbalanced Classification," *Machine Learning Mastery*, vol. 22, 2019.
- [23] N. V. Chawla, N. Japkowicz, and A. Kotcz, "Special issue on learning from imbalanced data sets," *ACM SIGKDD explorations newsletter*, vol. 6, no. 1, pp. 1-6, 2004.
- [24] Z. Bankó and J. Abonyi, "Correlation based dynamic time warping of multivariate time series," *Expert Systems with Applications*, vol. 39, no. 17, pp. 12814-12823, 2012.
- [25] H. Li, Y. Wang, H. Wang, and B. Zhou, "Multi-window based ensemble learning for classification of imbalanced streaming data," *World Wide Web*, vol. 20, no. 6, pp. 1507-1525, 2017.
- [26] T.-c. Fu, "A review on time series data mining," *Engineering Applications of Artificial Intelligence*, vol. 24, no. 1, pp. 164-181, 2011.
- [27] A. Arasu and G. S. Manku, "Approximate counts and quantiles over sliding windows," in *Proceedings of the twenty-third ACM SIGMOD-SIGACT-SIGART symposium on Principles of database systems*, 2004, pp. 286-296.
- [28] Z. Chen, Y. Liu, A. Valera-Medina, and F. Robinson, "A Multi-source Feature-level Fusion Approach for Predicting Strip Breakage in Cold Rolling," in *2020 IEEE 16th International Conference on Automation Science and Engineering (CASE)*, 2020, pp. 482-487: IEEE.
- [29] G. E. Batista, X. Wang, and E. J. Keogh, "A complexity-invariant distance measure for time series," in *Proceedings of the 2011 SIAM international conference on data mining*, 2011, pp. 699-710: SIAM.
- [30] M. Canizo, I. Triguero, A. Conde, and E. Onieva, "Multi-head CNN-RNN for multi-time series anomaly detection: An industrial case study," *Neurocomputing*, vol. 363, pp. 246-260, 2019.
- [31] T. Fawcett, "An introduction to ROC analysis," *Pattern recognition letters*, vol. 27, no. 8, pp. 861-874, 2006.