# A Hybrid Approach Combining Data-Driven and Signal-Processing-Based Methods for Fault Diagnosis of a Hydraulic Rock Drill

Hye Jun Oh[1], Jinoh Yoo[2], Sangkyung Lee[3], Minseok Chae[4], Jongmin Park[5], and Byeng D. Youn[6]*

[1,2,3,4,5,6]*Department of Mechanical Engineering, Seoul National University, Seoul, 08826, Republic of Korea*
*ojune575@snu.ac.kr*
*cosmbusi@snu.ac.kr*
*sangkyung25@snu.ac.kr*
*mschae1811@snu.ac.kr*
*20jmp02@naver.com*
*bdyoun@snu.ac.kr*

[6]*Institute of Advanced Machines and Design, Seoul National University, Seoul 08826, Republic of Korea*
*bdyoun@snu.ac.kr*

[6]*OnePredict Inc, Seoul, 08826, Republic of Korea*
*bdyoun@snu.ac.kr*

### ABSTRACT

This study presents a novel method for fault diagnosis of a hydrostatic rock drill. Hydraulic rock drills suffer from both domain discrepancy issues that arise due to their harsh working environment and indivisible difference. As a result, fault diagnosis is very challenging. To overcome these problems, we propose a novel diagnosis method that combines both data-driven and signal-process-based methods. In the proposed approach, data-driven methods are employed for overall fault classification, using domain adaptation, metric learning, and pseudo-label-based deep learning methods. Next, a signal-process-based method is used to diagnose the specific fault by generating a reference signal. Using the combined approach, the fault-diagnosis performance was 100%; the proposed method was able to perform well even in cases with domain discrepancy.

## 1. INTRODUCTION

A hydraulic rock drill (HRD) is a core component that is widely used in the mining, tunneling, and building industries. Such mechanical systems frequently work under harsh environmental conditions. A harsh environment is characterized by severe vibration, high humidity, and temperature variation; these issues cause diverse fault modes, including those of a missing seal and damage to a valve. As HRD faults cause huge economic and human costs due to work interruption, accurate fault diagnosis of HRDs is critical in real industrial fields.

To solve this problem, various fault-diagnosis studies have been conducted on HRDs. Previously proposed approaches can be largely divided into two methods: deep-learning-based methods and signal-processing-based methods. For signal-processing methods, first, features are extracted, and the pressure drop is studied to specifically capture the effect of accumulator performance (Erik, 2021). This feature is calculated as the slope of a straight line that is fit according to the physics equation. Prior research has confirmed that this feature is robust and allows accurate diagnosis of a pressure 'C' fault, which does not overlap with other faults in its distribution. This method has the advantage of being highly accurate and robust for a specific fault. However, a comprehensive diagnosis method that is useful for diagnosis of each fault mode remains to be developed, as it is a time-consuming and laborious task.

As a deep-learning-based method, a purely 1D-CNN deep-learning model that does not need feature engineering to diagnose the faults of a rock drill has been proposed (Senjoba, 2021). Compared to the signal-processing method, the deep-learning algorithm does not require expert knowledge and can perform well in multi-class fault diagnosis. However, its performance sharply decreases when the fault-diagnosis domain changes. For example, differences in the characteristics between different machines can cause changes


*Corresponding author: Byeng D. Youn (e-mail: bdyoun@snu.ac.kr)
https://doi.org/10.36001/IJPHM.2023.v14i1.3458

in the data distribution, which can make fault diagnosis difficult.

To overcome the limitations mentioned above, various studies have been conducted in industrial fields to attempt to identify more robust solutions. Existing approaches can be divided into three methods: domain adaptation, the pseudo-label technique, and metric learning. Developed a multi-adversarial learning strategy for different bearing datasets to obtain feature representations that are invariant to the multiple domain shifts. Song et al. introduced a retraining strategy with a weighted pseudo-label technique that is able to adapt to the test dataset; this approach was verified on bearing datasets.

With this context, the main contributions of this paper are as follows.

1) To the best of our knowledge, this research is the first study to increase the generalized performance of rock drill fault diagnosis in a situation studying many different devices.
2) The proposed approach improves generalized diagnostic performance using domain adaptation.
3) Ensemble deep-learning models are proposed to increase diagnostic reliability.
4) Faults that are difficult to diagnose through deep-learning are supplemented by a fault-specific signal processing-based method, leveraging the strengths of each approach.

## 2. Background

This section introduces the deep-learning-based and signal-processing-based techniques that are used in the proposed method. The concepts and usage of each technique are summarized.

### 2.1. Deep learning-based technique

The deep-learning-based techniques used in the proposed method consist of depth-wise convolution, domain adaptation, metric learning, and a pseudo-label technique.

#### 2.1.1. Depth-wise convolution

In contrast to conventional convolution operation, depth-wise convolution can extract unique features from each input channel. In particular, for a situation like this dataset that uses different sensors, a feature extractor for each sensor can extract sensor-specific information better than would be possible through general convolution. In addition, since convolution is carried out in a divided manner, the number of parameters can be reduced, as compared to general convolution. This may help prevent overfitting.

#### 2.1.2. Domain adaptation

In industrial sites, working conditions change constantly due to differences between users and devices; thus, there will always be discrepancies between the training and testing datasets. For this reason, it is impossible to diagnose a fault through the use of a general supervised learning technique. Many studies have been proposed to attempt to solve this problem by introducing domain adaptation (DA).

Li et al. (2021) proposed a deep CORAL method that was combined with CNN to deal with bearing fault diagnosis in several load torque conditions. Li et al.'s method extracted the fault feature using CNN and was more adaptive to diverse load torque conditions through its use of the deep CORAL approach. In other work, Mao et al. (2020) added loss functions (e.g., discriminative loss with maximum correlation entropy and a loss function that constrains the relatedness matrix L such that it is symmetric to the Domain Adversarial Neural Network (DANN)) to consider more information about the target domain. As a result, Mao et al.'s proposed method can effectively use DANN to extract domain-invariant features of a bearing under different working conditions.

For all DA methods described above, since learning is conducted with consideration of the target domain, the deep-learning model can be adapted even if data distribution occurs due to changes in working conditions or as a result of environmental changes.

#### 2.1.3. Metric learning

Metric learning means learning the distance between input data. In the process of learning embedding through similarity, there is an assumption that the process of identifying the similarity between data is to understand the input data well, that the purpose of the deep-learning model's learning is to understand the similarity between the data, and that this approach is an effective way to directly learn the similarity.

Wang et al. (2021) proposed a metric-learning model for imbalanced situations, for example, when the difference in the number of samples in each class is large. Wang et al.'s method enlarges the inter-class margin and compresses the intra-class angle distribution. This kind of metric-learning-based model can learn the relationship between each fault easily in conditions of imbalanced fault data.

#### 2.1.4. Pseudo label technique

The pseudo-label technique aims to help deep-learning models learn wider data distribution by imposing labels on the unlabeled data (Lee, 2013). These models can learn features from pseudo-labeled data by incorporating them into the training data. The general form of the objective function with pseudo-label loss integrated is as follows:

$$L_{tot} = \frac{1}{n}\sum_{m=1}^{n}\sum_{i=1}^{C} L_t(y_i^m, f_i^m) + \alpha\frac{1}{n'}\sum_{m=1}^{n'}\sum_{i=1}^{C} L_p(y'^m_i, f'^m_i)$$

where $L_t$ is cross-entropy loss based on the ground truth label, and $L_p$ is the loss based on the pseudo label. n and C refer to the number of data and classes. y and f are a ground truth label and a model prediction of the labeled data. y' and f' refers to the pseudo label and the model prediction of the unlabeled data, respectively. α is a constraint term that determines the extent to where which the pseudo label loss is used.

Recently, pseudo-label-based methods have been widely developed and have shown promising performance in diverse industrial fields. As unlabeled data is abundant in real industrial fields, in contrast to labeled data, the need for using the information from unlabeled data is significant. A bin pseudo label learning method was recently developed to reduce the distribution discrepancy between the source and target domains in bearing fault diagnosis (Yang et al, 2020). Other work introduced a pseudo-label-based retraining method, where the weights of each loss term are balanced; this approach was introduced to reduce the effect of noisy data to increase the generalization performance (Yan et al, 2019).

### 2.2. Signal processing-based technique

#### 2.2.1. Dynamic Time Warping

Dynamic Time Warping (DTW) is an optimization algorithm to align two signals (Sakoe and Chiba, 1978). The DTW approach extends or contracts the signal nonlinearly to find the warping path that minimizes the difference between two signals in time series. The warping path can be acquired by computing the cumulative distance matrix($\boldsymbol{\gamma}$) as follows.

$$\gamma(i, j) = d(a_i, b_j) + \min\{\gamma(i-1, j-1), \gamma(i-1, j), \gamma(i, j-1)\} \quad (1)$$

where $a$ and $b$ are the signals in time series and $d(\cdot,\cdot)$ is one of the distance metrics, including Euclidean distance and magnitude of difference.

The main advantage of DTW is that it makes two signals become similar shapes, although they have differences in terms of shifting and scaling (Cassisi, 2012). Further, DTW can also cope with different signal lengths. However, DTW can't be used as a metric because it does not follow the triangle inequality.

#### 2.2.2. Cross-correlation

Cross-correlation measures the similarity between one signal and a lagged signal. The cross-correlation definition is as follows.

$$\boldsymbol{r_{ab}(l)} = \frac{\sum_{i=1}^{n}(a_i - m_a)(b_{i-l} - m_b)}{\sqrt{\sum_{i=1}^{n}(a_i - m_a)^2}\sqrt{\sum_{i=1}^{n}(b_{i-l} - m_b)^2}} \quad (2)$$

where $\boldsymbol{m_a}$ and $\boldsymbol{m_b}$ are the means, and $\boldsymbol{l}$ is a time lag. The correlation $\boldsymbol{r_{ab}(l)}$ gives the degree of linear dependence between two signals, from -1 to 1[10].

## 3. Proposed Method

This section introduces the proposed method, which is divided into both deep-learning-based and signal processing-based methods, as shown in Figure 1. Figure 1 provides a flow chart of the proposed algorithm. The overall fault is diagnosed through a deep-learning algorithm, and samples that are difficult to diagnose through deep learning are diagnosed through a signal processing step.

$$Loss_{Total} = Loss_{Class} + Loss_{triplet} + Loss_{CORAL} + Loss_{pseudo} \quad (3)$$

### 3.1. Deep-learning-based method of the proposed approach

First, in order to be used for deep learning, the length of all input data must be constant. However, the length of the acquired data is different in most real-world settings because the operation time and users are different for each device. To address this issue, the slicing method was used. The data length was unified and cut to 557, the minimum length of the entire dataset. This strategy was largely implemented on three grounds. First, the accuracy when the data length was sliced from 0 to 556 was better than when sliced to 557 from the back and learned with resampled data. Second, for fault classes that are difficult to classify with deep learning, this was not a problem for the overall classification work because the data that contains the important information is in the 280-370 section.

After data preprocessing, each feature was extracted for three sensor data through a depth-wise-based encoder. The proposed depth-wise CNN is composed of multi-scale kernels. Since the area that can be convolved varies for each kernel size, the characteristics of the extracted features are also different. Consequently, this method of convolution is advantageous in the task of classifying many fault classes. During the training of the model, a slight difference in distribution may occur even in the same fault class, due to differences in the device; thus, metric learning was introduced to provide robust feature learning for multiple fault classifications. To this end, a special dataloader was used, and positive samples were randomly extracted from among classes, such as anchor samples, which can be a standard; negative samples randomly extracted from among other classes were learned at the same time. The special dataloader was calculated using the triplet loss function of the Euclidean-based loss function, combined with the margin so that the anchor and the positive could be close to each other and so that the anchor and negative could be pushed out of each other. Triplet loss was applied to a feature space consisting of 30 channels and the value was learned by backpropagating.

In addition, domain discrepancy occurs because each device is different. The coral loss was introduced in the proposed approach to reduce the effects of domain discrepancy. This domain adaptation approach was applied to the vector extracted through the global average pooling from among the classifier parts. Through this method, domain discrepancy can be reduced through a process of aligning the covariance of both the six training domains and the test domain.

Finally, a pseudo-labeling process was added to learn the information in the test dataset. The test dataset was evaluated with a model that learned from the training dataset, a pseudo label was given, and then the pseudo-labeled test dataset was retrained. This process updates the pseudo label and loss backpropagation for every 50th batch of the epoch.

The final best model can be obtained through this procedure. In the case of the total loss function, the model is learned by configuring it, as shown in the following equation.

The best model created was thus obtained, and to further compare the results, six models consisting of ablation models, additional models with attention, and serresnet were ensembled. Hard voting was used as the ensemble method, and the results obtained through this and the results obtained through the best model were compared. Most of the classes were well classified; however, the unmatched samples were extracted. In the independent seven dataset, classes 1 and 5 and classes 4 and 7 were partially different than the independent 8 dataset, so a specific signal-processing method was introduced in these four classes.

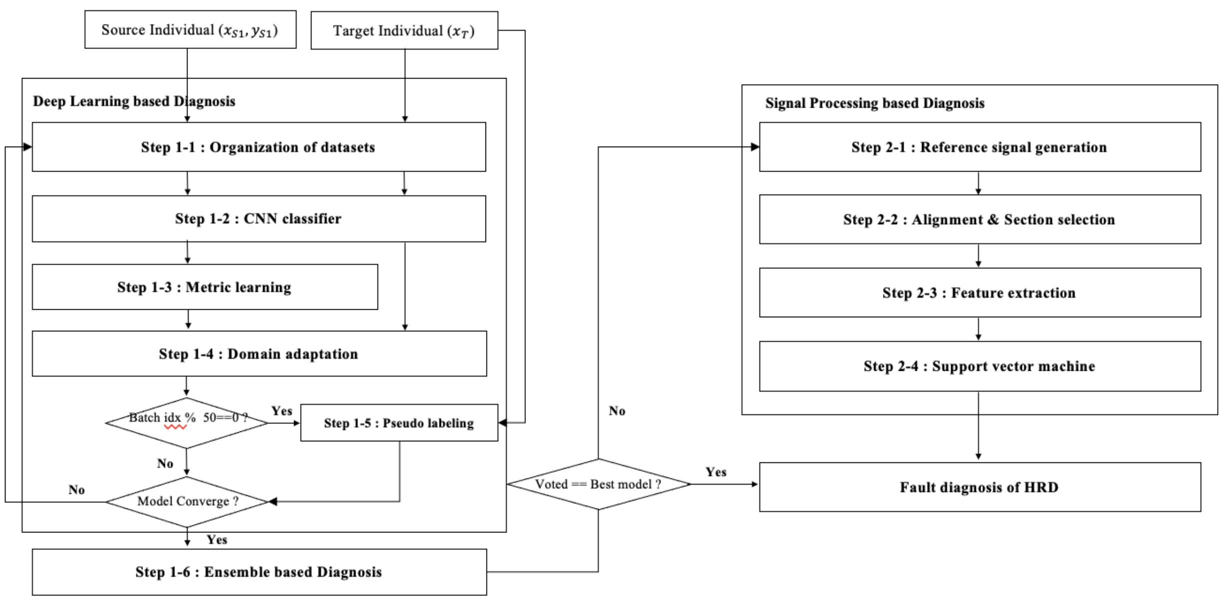



Figure 1. Flow chart of the proposed method

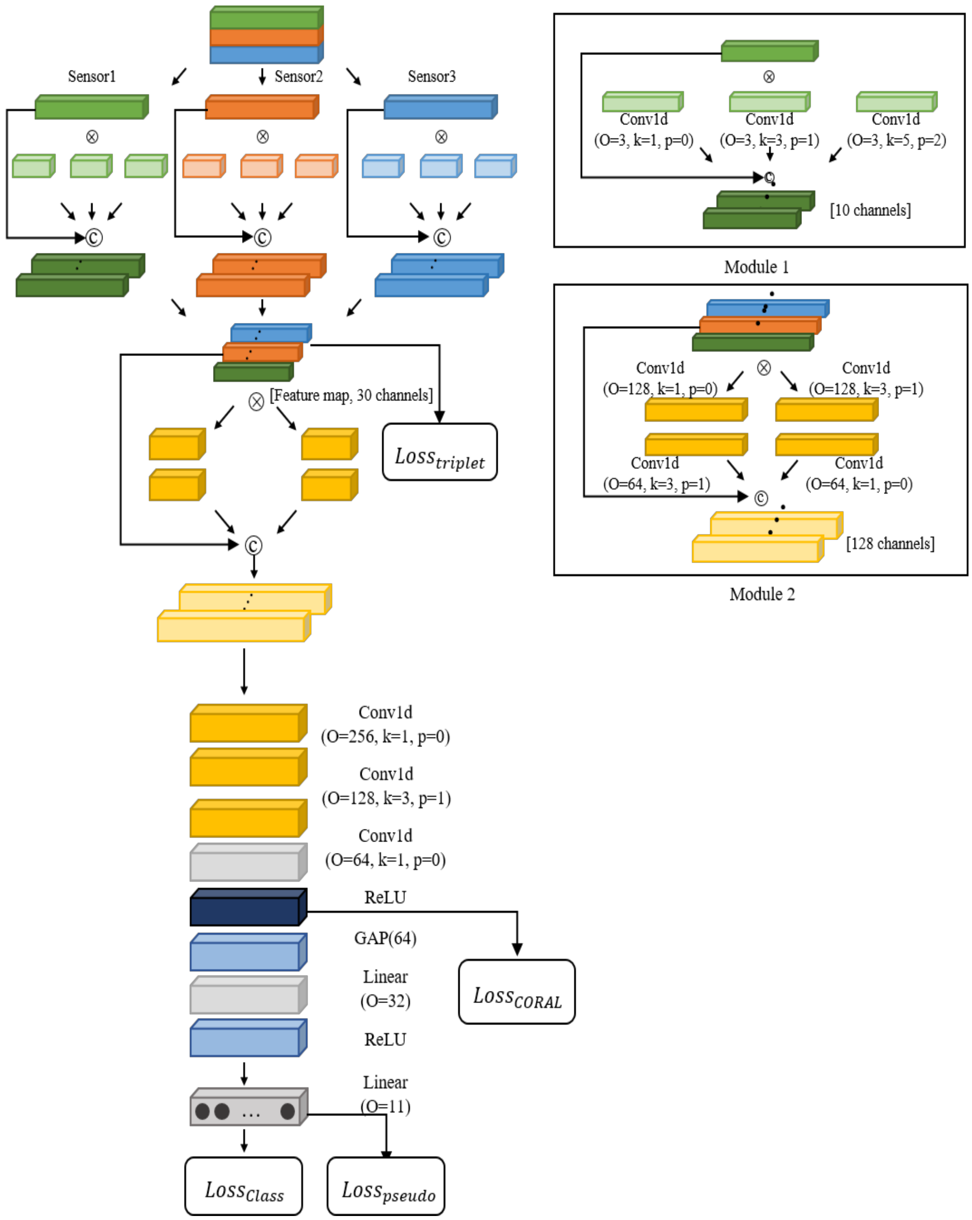



Figure 2. Framework of proposed method

### 3.2. Signal processing-based diagnosis

#### 3.2.1. Reference signal generation

Because all of the data, including the reference signals in the dataset, are not aligned and have different lengths, they need to be aligned. For that purpose, first, we generate three global reference signals (pin, pdin, pout) by using DTW and cross-correlation; these signals are used to align the whole dataset. The reference data are resampled into 748 samples. Then, using DTW, we align the reference signals from the same individual signal. To make a representative reference signal for each individual signal, we averaged the aligned signals. After acquiring the averaged reference signals from each individual signal, cross-correlation is used to align them. Finally, the signals that result from the cross-correlation are averaged to obtain three global reference signals that represent all of the individual reference signals.

#### 3.2.3. Feature Extraction

In the proposed method, three kinds of feature extraction methods are used in total. The statistical features are RMS, variance, and kurtosis, each of which is used to measure the energy, overall slope, and peakedness of the data.

- Statistical features from the aligned signals,
- Statistical features from the selected sections,
- Time lag value used in signal alignment.

For 1, the features are extracted to consider the statistical properties using the whole signal. For 2, which involves a deep investigation of the fault-specific features, selected sections are also used to extract the features. Finally, the time lags are used to consider the faults that are closely related to the time-domain delay.

Table 1. HRD fault modes

| Label | Letter | Description |
|---|---|---|
| 1 | NF | No-fault |
| 2 | T | Thicker drill steel |
| 3 | A | A-seal missing. Leakage from high pressure channel to control channel |
| 4 | B | B-seal missing. Leakage from the control channel to the return channel |
| 5 | R | Return accumulator, damaged |
| 6 | S | Longer drill steel |
| 7 | D | Damper orifice is larger than usual |
| 8 | Q | Low flow to the damper circuit |
| 9 | V | Valve damage. A small wear-flat on one of the valves lands |
| 10 | O | Orifice on the control line outlet is larger than usual. |
| 11 | C | The charge level in high pressure accumulator is low |

#### 3.2.2. Signal alignment and section selection

Before aligning, data are also resampled to have 748 samples per cycle. Then, cross-correlation is used to align the data by using the global reference signals, regardless of the individual signal. After aligning all the data, sections are selected at each pin, pdin, pout signal to capture the fault-specific features in this section. This section can be generalized into total data, due to the aligning procedure.

#### 3.2.4. Support Vector Machine

Support Vector Machine (SVM) is the supervised machine-learning model used for the classification problems. The hyperplane equation that divides the data points is as follows.

$$f(x) = w^T \mathrm{x} + \mathrm{b} = \sum_{j=1}^{M} w_j x_j + b = 0 \tag{4}$$

where the vector $\mathbf{w}$ is an M-dimensional vector and $\boldsymbol{b}$ is a bias term. To find the optimal hyperplane with slack variables, the following optimization problem is solved.

$$\text{minimize } \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{M}\xi_i \quad (5)$$

$$\text{s.t. } y_i(w^T \mathrm{x_i} + b) \geq 1 - \xi_i, \;\; i = 1, \cdots, M \quad (6)$$

where $\boldsymbol{\xi_i}$ measures the distance between the margin and the data $\mathbf{x_i}$ which are placed on the wrong side.

# 4. RESULTS

## 4.1. Problem Description

### 4.1.1. Problem definition: Hydraulic rock drill fault

In the 2022 PHM Data Challenge, we participated with the goal of classifying the fault of the Hydraulic Rock Drill. Our goal is to classify 11 classes, including ten faults of HDR and normal state. Labels 2 to 11 were assigned to 10 types of faults, and label one was assigned to the normal state. The ten faults are shown in Table 1.

### 4.1.2. Data description: Data size, number of individuals

The data used to solve the 11 classification problems are the training data obtained from five individuals and the validation data obtained from 1 individual. Using the data obtained from a total of 6 people, separate labels were assigned from individual 1 to individual 6. Individuals 1, 2, 4, 5, and 6 were used as the training data, and individual 3 was used as the validation data. There was a difference in the data length for each individual; the size of the data acquired from each individual is as shown in Table 2.

Also, the data acquired from each individual consists of data obtained from three types of sensors. Pressure measurement was performed on sensors attached to different locations, and the data acquired from each sensor were named “pin”, “pdin”, and “po”. “Pin” was obtained from the inlet fitting as the percussion pressure and “pdin” was measured in the outer chamber as the damper pressure. Finally, “po” is the pressure measured near the piston.

In summary, to train the classification model, data from six different individuals were used, with three types of data from each individual, giving a total of 18 types of data.

### 4.1.3. Scoring metric: Accuracy

To measure the performance of the algorithm, accuracy was used as a scoring metric. The algorithm was made to solve 11 classification problems, and the accuracy was calculated by comparing the predicted value with the actual value. To maximize the accuracy, the predicted fault mode was set to be the same as the actual fault mode.

## 4.2. Deep learning-based diagnosis

This section aims to show the diagnostic performance of the deep-learning method, from among the proposed methods. Figure 3 shows the overall diagnostic performance of each model. Three studies were conducted to show the performance of the proposed deep-learning algorithm. The first is a comparison with conventional methods for the classification tasks; the results of this performance comparison are summarized in Table 3. The second was the ablation study of the proposed method, which was implemented to understand the effect on the results of each step. The third is the extraction of unmatched samples by comparing the results with the ensemble-based method to determine which sample to use in the fault-specific signal process method.

In the first study, the performance of the proposed method was verified by comparing it with other existing methods. The first method is basic 1d CNN. This model evaluates the test dataset by learning only from information from individuals 1 to 6, without using any data from the target domain. The second model is a multi-kernel-based 1d CNN. This algorithm is advantageous for classification tasks of many classes because it can extract multiple pieces of information from the data using various kernels. For the third model, an algorithm was constructed by considering the information from the target domain. The third model used the maximum mean discrepancy (MMD), which reduces the distance between the two domains through a kernel-based approach. The fourth model used was the Deep CORAL model, which aligns the covariance of the source domain and

Table 2. Data descriptions

| Individual | Data type | Data Length |
|---|---|---|
| Individual 1 | Training | 7311 |
| Individual 2 | Training | 7867 |
| Individual 3 | **Validation** | 3184 |
| Individual 4 | Training | 7597 |
| Individual 5 | Training | 7977 |
| Individual 6 | Training | 3293 |

Table 3. Diagnosis Performance

| Methods | Model 1 | Model 2 | Model 3-1 | Model 3-2 | Model 4 | Proposed Method |
|---|---|---|---|---|---|---|
| Deep Learning Model | CNN | Multi-kernel CNN | Multi-kernel CNN | Multi-kernel CNN | Multi-kernel CNN | Multi-kernel CNN |
| Domain Adaptation | X | X | MMD | Deep CORAL | Deep CORAL | Deep CORAL |
| Metric Learning | X | X | X | X | Triplet | Triplet |
| Pseudo Label | X | X | X | X | X | O |
| Individual 7 Accuracy (%) | 98.79 | 98.88 | 99.22 | 99.86 | 99.87 | 100 |
| Individual 8 Accuracy (%) | 86.14 | 89.53 | 96.14 | 99.35 | 99.46 | 100 |

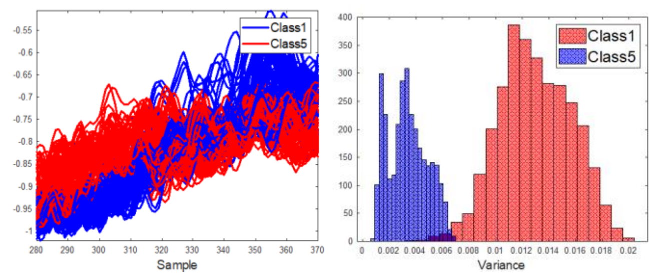
the target domain. The fifth model used the metric learning to consider the relationship between each fault condition.

Compared to these methods, the proposed method performed overwhelmingly well for both the individual 7 dataset and the individual 8 dataset. This method was able to show excellent performance because it reduced the domain discrepancy, while considering the relationship between each class. In addition, as shown in Table 3, it was confirmed that the performance increased dramatically when domain adaptation was used. Although the proposed approach shows sufficiently high performance already, it is possible to record 100% performance because the metric learning enables more robust feature learning and uses a pseudo-label approach to more actively utilize information in the target domain.

### 4.3. Signal processing-based diagnosis performance

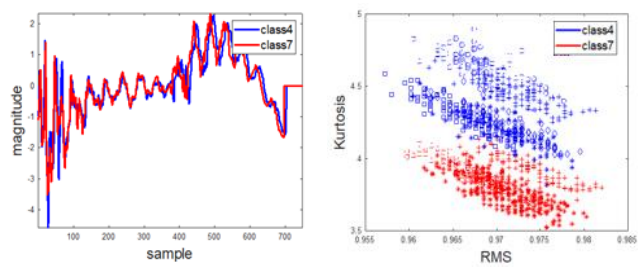
As mentioned in the introduction, the signal-processing-based method was used to classify the confused fault modes, which are 1 & 5, 4 & 7. For these classes, 280~370, 365~415 from the Pout signal and 115~195, 270~330 from the Pdin signal were chosen as the sections. Then all features were used in the SVM to classify the fault modes. To show how the features work, the following figures are provided. Figure 3(a) shows the slope difference clearly between class 1 and class 5 for the 280~370 samples in the Pout signal. Further, Figure 3(b) shows a clear difference in the variance distribution for classes 1 & 5. In Figure 4(a), it is hard to see the difference between those two Pin signals at a glance. However, Figure 4(b) shows the distributions separately expressed by RMS and kurtosis, which used aligned signals for classes 4 & 7.

After extracting all features, SVM was trained using the training data. In this method, all the datasets from individuals 1~6 and the confident data (not-confused) from individuals 7 and 8 was used to train the SVM model.




(a) (b)

Figure 3. (a) Slope difference between classes 1 & 5 at the 280 to 370 samples in the Pout signal, (b) Clearly separated distributions of variance for the selected section




(a) (b)

Figure 4. (a) The aligned Pin signals (b) Clearly separated distributions with RMS and kurtosis features

Thus, the signal-processing-based method was able to make a perfect decision as to which fault mode was correct for the confusing data.

### 5. Conclusion

This paper outlines our proposed method to diagnose the failure of a hydraulic rock drill. To solve 11 classification problems, we proposed a hybrid of both deep-learning and signal-processing techniques to maximize the overall algorithm's performance.

The first step for the deep-learning method is feature extraction using a depth-wise-based encoder in the deep-learning model. In this case, a Euclidean-based loss function using triplet loss was used as the loss function. Also, for this data, since there is a domain discrepancy between the individuals' datasets, a coral loss is used to reduce the domain discrepancy. Finally, through a pseudo-labeling process, we used the test dataset information for algorithm learning to build a deep-learning model.

We also proposed to increase the accuracy by applying an ensemble-based voting mechanism to the failure mode predicted by the deep-learning model. Via the voting mechanism, we extracted data that predicted a different fault mode from the best model and applied a signal-processing-based diagnostic model to the data. In the signal-processing-based diagnostic model, we checked the failure mode by comparing it with the aligned dataset through DTW and cross-correlation. Further, if there were data that had not been clearly classified, the failure mode for those data was determined through the SVM algorithm, using statistical features.

The results show that our hybrid algorithm achieves high accuracy of 100% in HRD fault diagnosis. In future work, hyperparameter optimization will be performed to make the hybrid algorithm robust for particular real-world circumstances. Moreover, a resampling method will be implemented to avoid losing information.

## Acknowledgement

This research was financially supported by the Ministry of Trade, Industry and Energy (MOTIE) and Korea Institute for Advancement of Technology (KIAT) through the International Cooperative R&D Program (P0011923) and HRD Program for Industrial Innovation (P0008691).

## References

Jakobsson, E., Frisk, E., Krysander, M., & Pettersson, R. (2021). Fault Identification in Hydraulic Rock Drills from Indirect Measurement During Operation. *IFAC-PapersOnLine*, 54(11), 73-78.

Senjoba, L., Sasaki, J., Kosugi, Y., Toriya, H., Hisada, M., & Kawamura, Y. (2021). One-Dimensional Convolutional Neural Network for Drill Bit Failure Detection in Rotary Percussion Drilling. *Mining*, 1(3), 297-314.

Li, X., Zhang, Z., Gao, L., & Wen, L. (2021). A New Semi-Supervised Fault Diagnosis Method via Deep CORAL and Transfer Component Analysis. *IEEE Transactions on Emerging Topics in Computational Intelligence*, 6(3), 690-699.

Mao, W., Liu, Y., Ding, L., Safian, A., & Liang, X. (2020). A new structured domain adversarial neural network for transfer fault diagnosis of rolling bearings under different working conditions. *IEEE Transactions on Instrumentation and Measurement*, 70, 1-13.

Wang, C., Xin, C., & Xu, Z. (2021). A novel deep metric learning model for imbalanced fault diagnosis and toward open-set classification. *Knowledge-Based Systems*, 220, 106925.

Lee, D. H. (2013, June). Pseudo-label: The simple and efficient semi-supervised learning method for deep neural networks. *In Workshop on challenges in representation learning, ICML* (Vol. 3, No. 2, p. 896).

Yang, B., Lei, Y., Jia, F., & Xing, S. (2019). An intelligent fault diagnosis approach based on transfer learning from laboratory bearings to locomotive bearings. *Mechanical Systems and Signal Processing*, 122, 692-706.

Zhang, K., Chen, J., Zhang, T., He, S., Pan, T., & Zhou, Z. (2020). Intelligent fault diagnosis of mechanical equipment under varying working condition via iterative matching network augmented with selective Signal reuse strategy. *Journal of Manufacturing Systems*, 57, 400-415.

Sakoe, H., & Chiba, S. (1978). Dynamic programming algorithm optimization for spoken word recognition. *IEEE transactions on acoustics, speech, and signal processing*, *26*(1), 43-49.

Cassisi, C., Montalto, P., Aliotta, M., Cannata, A., & Pulvirenti, A. (2012). Similarity measures and dimensionality reduction techniques for time series data mining. *Advances in data mining knowledge discovery and applications*, 71-96.

Wan, L., Li, Y., Chen, K., Gong, K., & Li, C. (2022). A novel deep convolution multi-adversarial domain adaptation model for rolling bearing fault diagnosis. Measurement, 191, 110752.

Song, Y., Li, Y., Jia, L., & Qiu, M. (2019). Retraining strategy-based domain adaption network for intelligent fault diagnosis. IEEE Transactions on Industrial Informatics, 16(9), 6163-6171.

**BIOGRAPHIES**

**Hye Jun Oh** received the B.S. degree in mechanical engineering from Sungkyunkwan University, Seoul, Republic of Korea, in 2022. He is currently pursuing the Ph.D. degree in the Department of Mechanical Engineering at Seoul National University, Seoul, Republic of Korea. His current research topics include prognostics and health management for electric machines using deep-learning approaches.

**Jinoh Yoo** received the B.S. degree in mechanical engineering from Yonsei University, Seoul, Republic of Korea, in 2021. He is currently pursuing the Ph.D. degree in the Department of Mechanical Engineering at Seoul National University, Seoul, Republic of Korea. His current research interests include prognostics and health management for gearboxes and industrial robots using signal processing.

**Sangkyung Lee** received the B.S. degree in mechanical engineering from Yonsei University, Seoul, Republic of Korea, in 2022. He is currently pursuing the Ph.D. degree in the Department of Mechanical Engineering at Seoul National University, Seoul, Republic of Korea. His current research interests include prognostics and health management for electric drive systems, and deep learning algorithms.

**Minseok Chae** received the B.S. degree in mechanical engineering from Chung-Ang University, Seoul, Republic of Korea, in 2019. Subsequently, he worked as an engineer to solve the failure of semiconductor equipment at SK Hynix System IC Inc. from 2019.1 to 2020.8. He is currently pursuing the Ph.D. degree in the Department of Mechanical Engineering at Seoul National University, Seoul, Republic of Korea. His current research topics include prognostics and health management for electromechanical drive systems using deep-learning approaches.

**Jongmin Park** received the B.S. degree from Seoul National University, Seoul, South Korea, in 2017, where he is currently pursuing the Ph.D. degree in mechanical and aerospace engineering. His current research interests include prognostics and health management for rolling element bearings and gas insulated switchgear. He received an award as the PHM Asia Pacific Data Challenge Winner, in 2017 and 2019

**Byeng D. Youn** is the Full Professor of Mechanical Engineering at Seoul National University (SNU) and the founder and CEO of OnePredict, Inc. (onepredict.com). Prof. Youn received the B.S. degree in mechanical engineering from Inha University, Incheon, South Korea, in 1996, the M.S. degree in mechanical engineering from Korea Advanced Institute of Science & Technology, Daejeon, South Korea, in 1998, and the Ph.D. degree in mechanical engineering from the University of Iowa, Iowa City, IA, USA, in 2001. His current research includes prognostics and health management (PHM), engineering design under uncertainty, and energy harvester design. His dedication and efforts in research have garnered substantive peer recognition, resulting in many notable awards including the Commendation of Prime Minister (2019), the Shin Yang Academic Award from Seoul National University (2017), the IEEE PHM Competition Winner (2014), the PHM Society Data Challenge Winners (2014, 2015, 2017, 2019), the Best Paper Award of ICMR2019 (2019), the PHMAP 2019 PHM Data Challenge Competition 3rd Place Winners (2019), the Front Cover and Featured Article of Journal of Applied Physics (2020), the Industrial AI Hackathon Factory Hack Korea 2021 Winners (2020), among others.