

Unilateral Alignment: An interpretable machine learning method for geophysical logs calibration

Wenting Zhang^a, Jichen Wang^b, Kun Li^c, Haining Liu^e, Yu Kang^{a,c,d}, Yuping Wu^d, Wenjun Lv^{a,c,*}

^a Department of Automation, University of Science and Technology of China, Hefei, 230027, China

^b College of Control Science and Engineering, China University of Petroleum, Qingdao, 266580, China

^c Institute of Advanced Technology, University of Science and Technology of China, Hefei, 230031, China

^d Institute of Artificial Intelligence, Hefei Comprehensive National Science Center, Hefei, 230088, China

^e Shengli Geophysical Research Institute, SINOPEC Group, Dongying, 257022, China

ARTICLE INFO

Keywords:

Interpretable machine learning
Geophysical logs calibration
Data distribution discrepancy

ABSTRACT

Most of the existing machine learning studies in logs interpretation do not consider the data distribution discrepancy issue, so the trained model cannot well generalize to the unseen data without calibrating the logs. In this paper, we formulated the geophysical logs calibration problem and give its statistical explanation, and then exhibited an interpretable machine learning method, i.e., Unilateral Alignment, which could align the logs from one well to another without losing the physical meanings. The involved UA method is an unsupervised feature domain adaptation method, so it does not rely on any labels from cores. The experiments in 3 wells and 6 tasks showed the effectiveness and interpretability from multiple views.

1. Introduction

Artificial intelligence (AI) is increasingly becoming an important scientific and technological tool in solving geosciences issues such as earthquakes detection, spatial prediction, stratigraphic correlation, seismic processing and interpretation, etc. (Magrini et al., 2020; Fouedjio, 2020; Xu et al., 2022; Zhou et al., 2020; Birnie et al., 2021). It has reported that the rapidly evolving field of machine learning (ML), a popular way to realize AI, will play a key role in geosciences (Bergen et al., 2019). However, due to the black-box issue existing in many ML models, we cannot know the way how an ML model works, so cannot assess its reliability and have enough confidence to apply it in production. We believe that the interpretability of ML will play an extremely important role in geosciences in which there are too many risk-aware applications, e.g., exploration.

The geophysical logs interpretation means the prediction of borehole geology information (e.g., lithofacies, porosity) according the logs, which is a fundamental and worthy work to understand the undersurface earth. The recent years have witnessed the development of the applications of machine learning in logs interpretation. For example, a semi-supervised model named Laplacian support vector machine is proposed to solve the

problems of scarce labels (Li et al., 2020). The smoothnesses implying in the feature space and depth are utilized to propagate the labels from labelled samples to unlabelled ones, therefore increasing the classification accuracy. Although the Laplacian works in semi-supervised learning, the setting of Laplacian is still empirical so that the introduction of unlabelled data might result in a worse performance compared with just using labelled data. Hence, an ensemble mechanism is used to combine some candidate Laplacians and calculate an optimal one to guarantee the safety of using semi-supervision (Li et al., 2021). More similar work could be found in the literatures (Dunham et al., 2020; Imamverdiyev and Sukhostat, 2019; Zhu et al., 2018).

Although there has been a great deal of researches in this area, most of them work under the assumption of independent and identically distribution (iid), i.e., the data from training wells should have similar distribution with those data from testing wells. However, such an assumption barely hold due to the differences in logging equipments, borehole conditions, etc. As illustrated in Fig. 1c, there are a significant distribution discrepancy between the training dataset and testing dataset, so iid does not hold anymore. As shown in Fig. 1a and b, the non-iid problem causes an accuracy decrease in interpreting the logs from another wells by the trained classifier. Both supervised and semi-

* Corresponding author. Department of Automation, University of Science and Technology of China, Hefei, 230027, China.

E-mail address: wlv@ustc.edu.cn (W. Lv).

<https://doi.org/10.1016/j.aig.2022.02.006>

Received 27 December 2021; Received in revised form 27 February 2022; Accepted 27 February 2022

Available online 14 March 2022

2666-5441/© 2022 The Authors. Publishing Services by Elsevier B.V. on behalf of KeAi Communications Co. Ltd. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

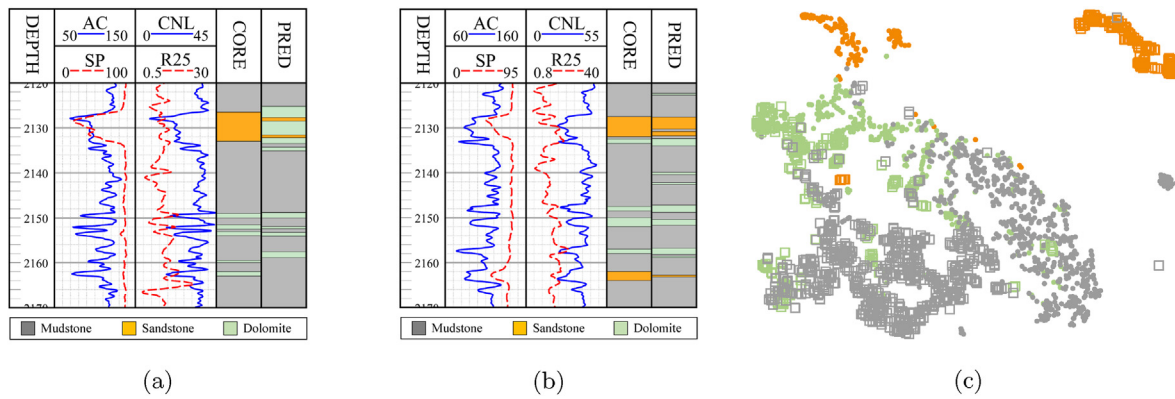


Fig. 1. Illustration of data distribution discrepancy. (a) Logs, cores (i.e., labels), and predictions on Well A using the model trained on Well B; (b) Logs, cores, predictions on Well B using the model trained on Well A; (c) t-SNE (van der Maaten and Hinton, 2008) visualization of data distribution for Well A and Well B. Gray, yellow, and green indicate mudstone, sandstone, and dolomite, respectively. Dots and boxes indicate the data from Well A and Well B, respectively.

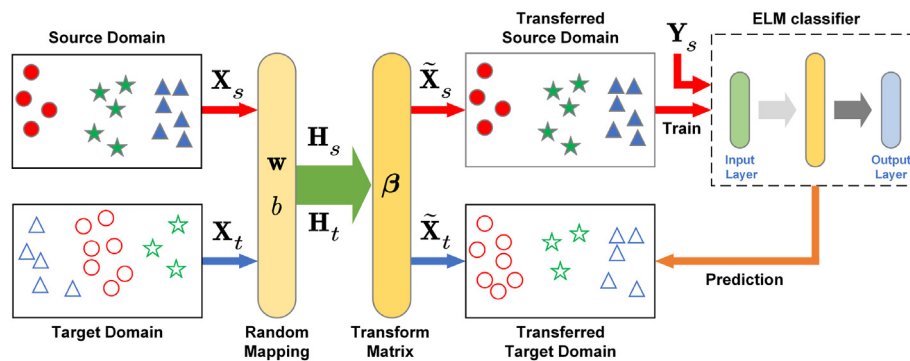


Fig. 2. Illustration of the UA framework.

Table 1

Dataset description.

Wells ^b	Lithology ^a			
	Mu	Si	Co	Sum
A	955	512	113	1580
B	514	396	108	1018
C	735	220	100	1055

^a Mu, mudstone; Si, siltstone; Co, conglomeratic sandstone.

^b Logging curves of each well: AC, acoustic log; CAL, caliper log; COND, conductivity log; GR, gamma ray log; R25, 2.5 m bottom gradient resistivity; and SP, spontaneous potential log resistivity.

Table 2

Macro-average F1-score (%) on 6 tasks.

Tasks	Methods		
	S.O. ^a	T.O. ^b	UA
B → A	82.7	98.5	95.9
C → A	80.0	99.3	92.7
A → B	82.9	99.9	97.2
C → B	79.7	99.5	93.5
A → C	92.3	99.4	96.1
B → C	73.3	98.4	93.7

^a Weighted ELM classifier trained with source-domain examples.

^b Weighted ELM classifier trained with target-domain examples.

supervised classification methods cannot handle such a situation, so some pre-processing work should be done in advance of interpretation. In

addition, data mining and processing techniques can be coupled with machine learning algorithms to solve the problem of data imbalance (Long et al., 2016; Zhong et al., 2020). Filtering techniques help us to remove some of the random non-stratigraphic responses (Lv et al., 2018; Ruckebusch, 1983), but do not have a significant effect on the distribution alignment. For such a geophysical logs calibration issue, we can resort to the Domain Adaptation (DA) methods which are very popular in machine learning.

As a transfer learning method, DA tries to find the relevance between two datasets and then adopts the knowledge from one domain to another. DA could works in an unsupervised manner which means that we only need the samples from two datasets. DA generally falls into three categories: model-, sample-, and feature-based methods (Yang et al., 2020). The sample-based DA re-weights the samples overlapped in the feature space during the training phase to yield a more robust model in both domain, but it could only handle the small distribution discrepancy issue (Chang et al., 2021b; Chang et al., 2022). The model-based DA re-adjust the source-domain classifier to fit the target domain (Liu et al., 2020). For a neural network model, usually the last few layer are allowed to adjust, so the transferability is not strong enough. The feature-based DA (fDA) tries to map the samples to a feature space with higher dimension such that two domains align (Li et al., 2018). As we know that a deep neural network possesses multiple layers for feature extraction, so the fDA usually has more feasibility. Most fDA methods map both domains, so the physical meanings of the original features would loss (Chang et al., 2021a). Instead, we could just align one domain to another in the original feature space, therefore maintaining the physical meanings (Chen et al., 2018; Wu et al., 2022). Such a fDA method could be named Unilateral Alignment (UA) which might be an excellent tool for geophysical logs calibration to our opinions.

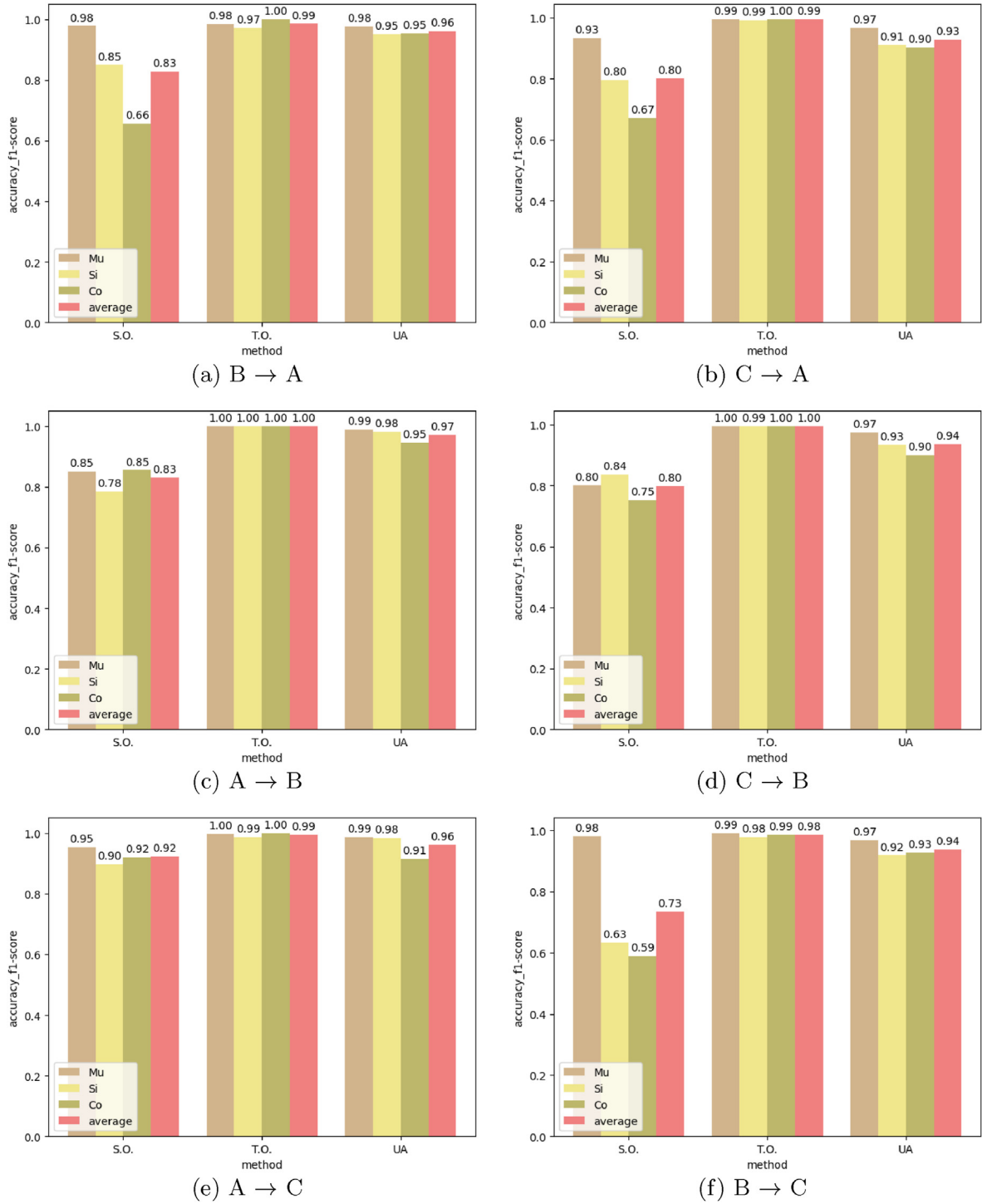


Fig. 3. F1-score of each class in six tasks.

In this paper, we would exhibit an interpretable machine learning method, i.e., UA, which could align the logs from one well to another without losing the physical meanings. In addition, the involved UA method is an unsupervised FDA method, so it does not rely on any labels from cores. The contributions are threefold. First, we formulate the geophysical logs calibration problem and give its statistical explanation. Second, we present the derivation of UA and illustrate its application in logs calibration. Third, we conduct massive experiments in 3 wells and 6 tasks to show the effectiveness and interpretability from multiple views. The above three contributions correspond to Sec. 2, 3, and 4, respectively.

2. Formulation of geophysical logs calibration problem

In the logs interpretation problem (e.g., lithofacies identification), the well logging feature vector (also named instance or sample) of d logs at a certain depth is denoted by $\mathbf{x}_i = [x_{i1}; x_{i2}; \dots; x_{id}] \in \mathbb{R}^d$, including acoustic log (AC), Gamma ray log (GR), caliper log (CAL), etc. The lithofacies corresponding to the instance \mathbf{x}_i is denoted by

$$\mathbf{y}_i = [\underbrace{0, \dots, 0}_{k_0-k}, 1, \underbrace{0, \dots, 0}_k],$$

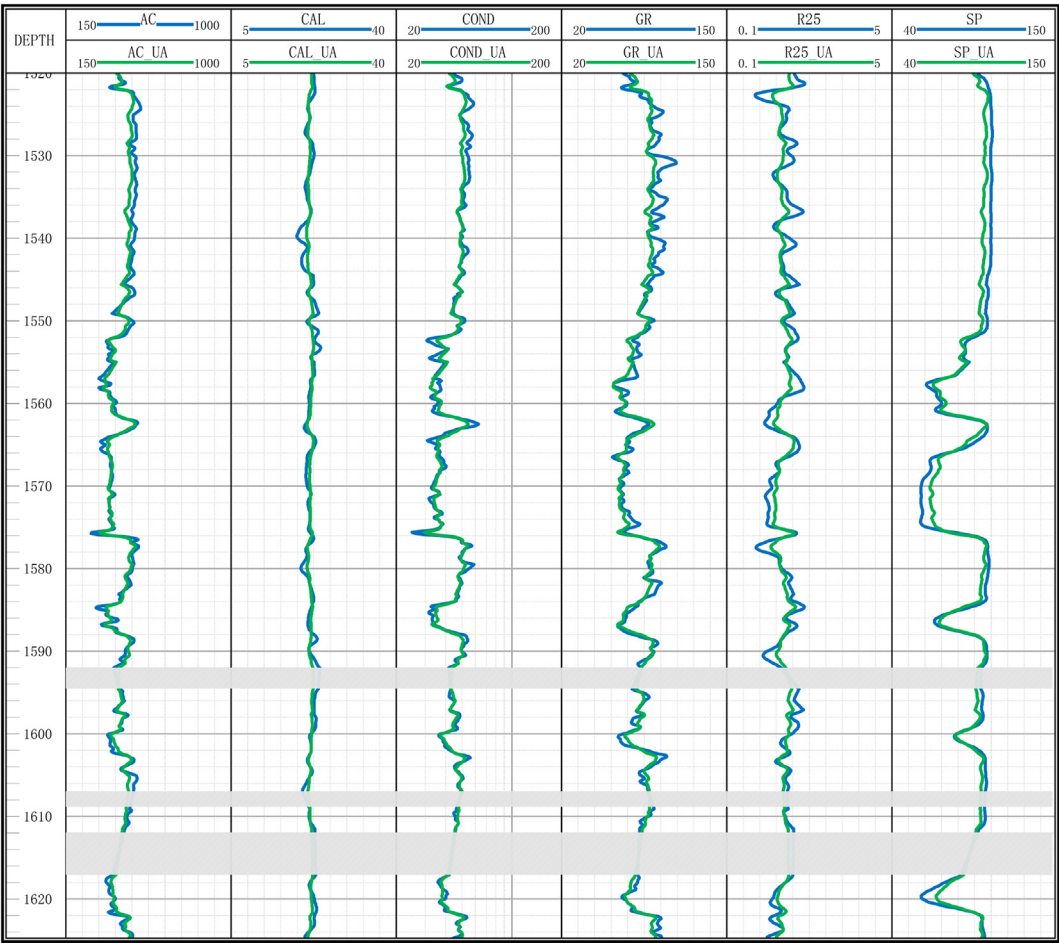


Fig. 4. Visualization of Well A logs in the task A → B before and after calibration.

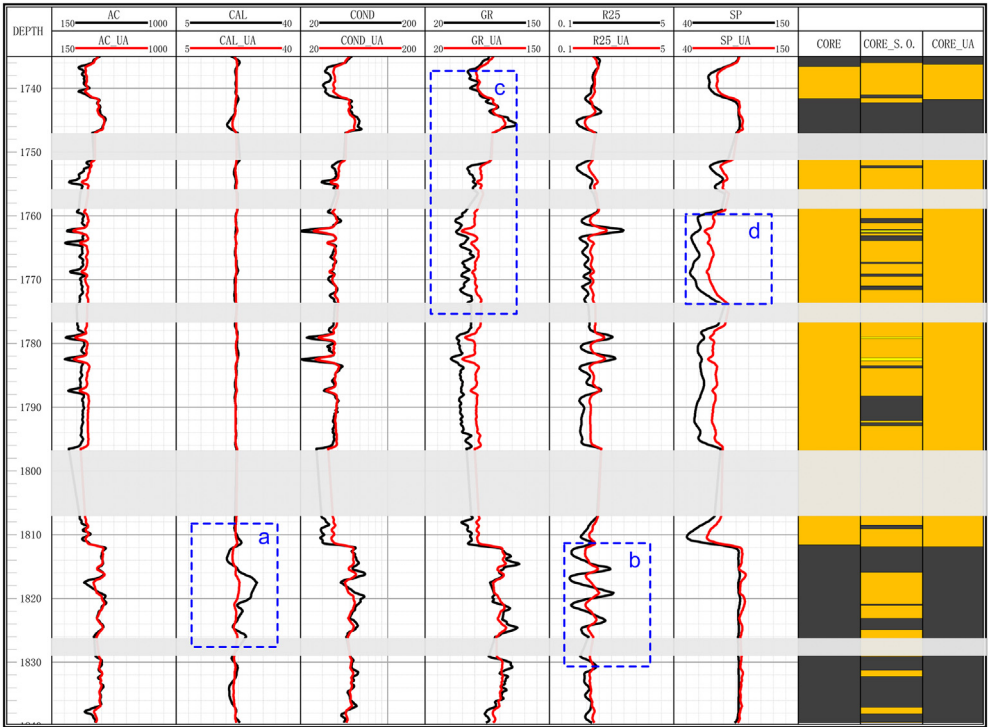


Fig. 5. Visualization of Well B logs and the corresponding predicted lithologies in the task A → B before and after calibration.

Table 3

The mean values of each log on the task A → B.

Domain	Mean					
	AC	CAL	COND	GR	R25	SP
Original Source	429.50	21.81	43.50	81.41	1.80	90.69
Aligned Source	429.07	21.18	43.45	81.32	1.79	90.62
Original Target	417.17	22.25	44.31	74.22	1.68	87.84
Aligned Target	429.07	21.18	43.45	81.32	1.79	90.62

Table 4

The mean values of each log on the task A → C.

Domain	Mean					
	AC	CAL	COND	GR	R25	SP
Original Source	429.50	21.81	43.50	81.41	1.80	90.69
Aligned Source	429.50	21.81	43.50	81.41	1.80	90.69
Original Target	438.92	22.03	44.96	74.37	1.83	96.41
Aligned Target	437.15	21.98	43.66	79.69	1.82	95.44

when \mathbf{x}_i belongs to the k -th ($1 \leq k \leq k_0$) lithofacies type of the total k_0 types. n examples $(\mathbf{x}_i, \mathbf{y}_i)$ along the depth constitute a labelled dataset $D = \{(\mathbf{x}_i, \mathbf{y}_i)\}_{i=1}^n$ from the well.

In most cases, the source-domain instances and the target-domain ones have different probability distributions. With a logs calibration, the accuracy of lithofacies identification might decrease when predicting the target-domain instances by the model trained on source-domain examples. In this case, the expert-interpreted dataset is denoted as the source-domain dataset $D_s = \{(\mathbf{x}_i^s, \mathbf{y}_i^s)\}_{i=1}^{n_s}$, and the dataset of an uninterpreted well is called target-domain dataset $D_t = \{(\mathbf{x}_j^t, \mathbf{y}_j^t)\}_{j=1}^{n_t}$, where n_s, n_t represent the number of the instances in D_s, D_t , respectively. Besides, these instances of both domains belong to the same feature space \mathcal{X} , which means that $\mathbf{x}^s, \mathbf{x}^t \in \mathcal{X}$. The label space \mathcal{Y} of D_s is also identical with D_t 's, i.e., $\mathbf{y}^s, \mathbf{y}^t \in \mathcal{Y}$.

Since the datasets of two domains are sampled from respective joint distributions, the discrepancies of probability distributions between two domains can be seen as the discrepancies of two joint distributions i.e., $P_s(\mathbf{x}^s, \mathbf{y}^s) \neq P_t(\mathbf{x}^t, \mathbf{y}^t)$ (Vapnik, 1999). The joint distributions of D_s and D_t are rewritten as $P_s(\mathbf{x}^s, \mathbf{y}^s) = P_s(\mathbf{x}^s)P_s(\mathbf{y}^s|\mathbf{x}^s)$, $P_t(\mathbf{x}^t, \mathbf{y}^t) = P_t(\mathbf{x}^t)P_t(\mathbf{y}^t|\mathbf{x}^t)$,

respectively, where $P_s(\cdot), P_t(\cdot)$ are marginal probability distributions, and $P_s(\cdot|\cdot), P_t(\cdot|\cdot)$ are conditional ones. In general, the conditional probability distributions are shared across domains (Stojanov et al., 2021). Under the assumption of $P_s(\mathbf{y}^s|\mathbf{x}^s) = P_t(\mathbf{y}^t|\mathbf{x}^t)$, the problem turns to tackle the inequality of marginal probability distributions, i.e., $P_s(\mathbf{x}^s) \neq P_t(\mathbf{x}^t)$.

In order to improve the performance in the case of probability discrepancy, we propose a calibration method to learn a unilateral calibrating function f compelling $P_s(f(\mathbf{x}^s)) \approx P_t(f(\mathbf{x}^t))$ and consequently $P_s(f(\mathbf{x}^s), \mathbf{y}^s) \approx P_t(f(\mathbf{x}^t), \mathbf{y}^t)$. In this way, the classifier trained on D_s would have a good performance in the interpretation task on a target-domain well.

3. Unilateral Alignment for geophysical logs calibration

We first give some preliminaries in Sec. 3.1 and 3.2, and then yields the Unilateral Alignment (UA) algorithm in Sec. 3.3.

3.1. Extreme Learning Machine (ELM)

Extreme Learning Machine (ELM) can be used to address classification and regression problems, which is proposed by G.-B. Huang (Huang et al., 2004, 2011). Its model training is efficient compared with some traditional learning algorithms, meantime completing regression and multi-class classification tasks effectively. For example, considering a classification task with training dataset $\{\mathbf{X}, \mathbf{Y}\} = \{(\mathbf{x}_i, \mathbf{y}_i)\}_{i=1}^n$, ELM corresponding to the training dataset with L hidden layer nodes is described as

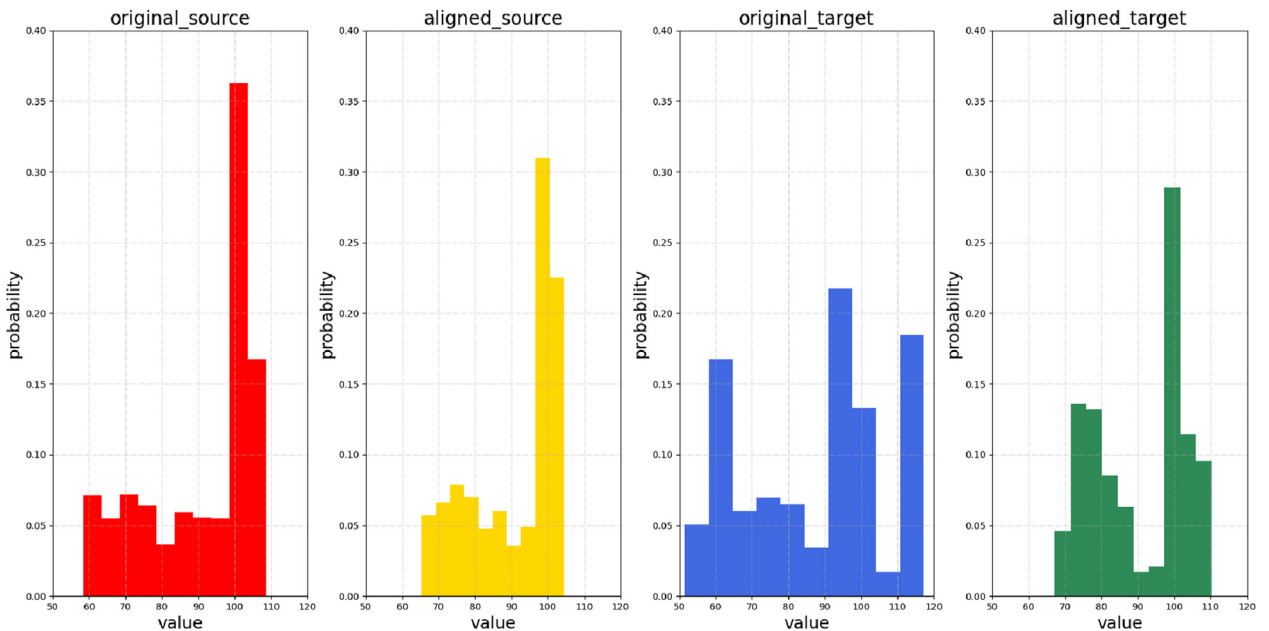
$$\sum_{i=1}^L \beta_i g(\mathbf{w}_i^\top \mathbf{x}_j + b_i) = \mathbf{o}_j, j = 1, \dots, n, \quad (1)$$

where $g(\cdot)$ is the nonlinear piecewise continuous activation function (e.g., Sigmoid function), $\mathbf{w}_i = [w_{i1}; w_{i2}; \dots; w_{id}] \in \mathbb{R}^d$ and b_i are the randomly-generated weight vector and bias of the i -th hidden layer node, β_i is the output weight of the i -th hidden layer node. \mathbf{o}_j is the prediction of the j -th instance.

To minimize the predicted error, the error function can be defined as

$$\sum_{j=1}^n \|\mathbf{o}_j - \mathbf{y}_j\| = 0, \quad (2)$$

where \mathbf{y}_j denotes the j -th label among the n instances. Combining (1) and

**Fig. 6.** Frequency distribution histogram of SP log in the task A → B.

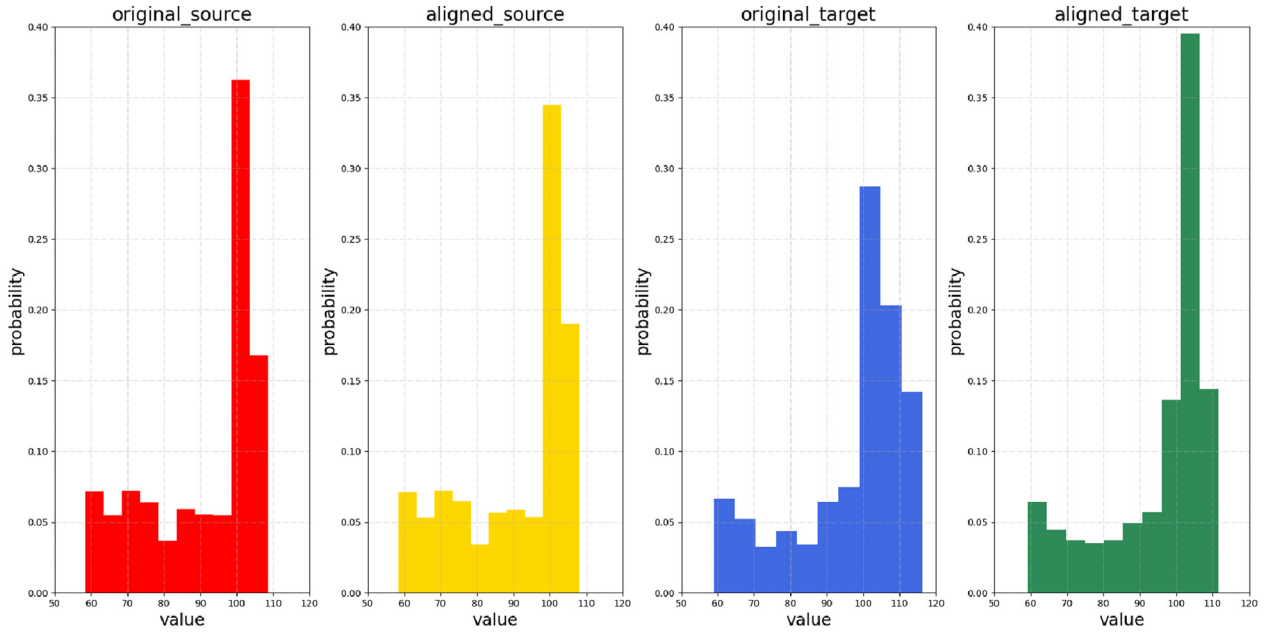


Fig. 7. Frequency distribution histogram of SP log in the task A → C.

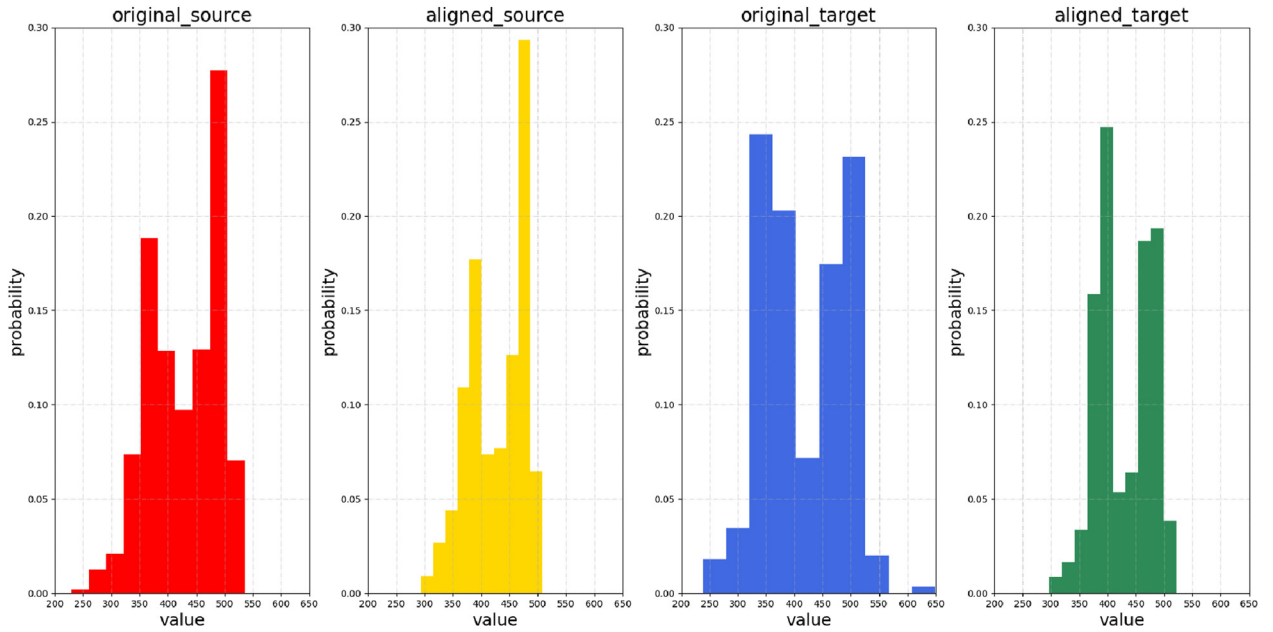


Fig. 8. Frequency distribution histogram of AC log in the task A → B.

(2) yields $\mathbf{H}\boldsymbol{\beta} = \mathbf{Y}$, where $\mathbf{H} = [\mathbf{h}(\mathbf{x}_1); \mathbf{h}(\mathbf{x}_2); \dots; \mathbf{h}(\mathbf{x}_n)] \in \mathbb{R}^{n \times L}$ represents the mapped instance matrix, and $\boldsymbol{\beta} \in \mathbb{R}^{L \times k_0}$ denotes the output weight matrix, $\mathbf{h}(\mathbf{x}_i) = [g(\mathbf{x}_i \mathbf{w}_1 + b_1), g(\mathbf{x}_i \mathbf{w}_2 + b_2), \dots, g(\mathbf{x}_i \mathbf{w}_L + b_L)] \in \mathbb{R}^{1 \times L}$ is the i -th mapped instance. Hence, the loss function is

$$\min_{\boldsymbol{\beta}} \|\mathbf{H}\boldsymbol{\beta} - \mathbf{Y}\|^2, \quad (3)$$

where $\boldsymbol{\beta}$ could be calculated by the Moore-Penrose (MP) generalized inverse, that is,

$$\boldsymbol{\beta}^* = \mathbf{H}^\dagger \mathbf{Y}, \quad (4)$$

where \mathbf{H}^\dagger is the MP generalized inverse matrix of \mathbf{H} . To avoid the over-fitting issue, (3) could be augmented as

$$\min_{\boldsymbol{\beta}} \frac{1}{2} \|\boldsymbol{\beta}\|^2 + \frac{C}{2} \|\mathbf{H}\boldsymbol{\beta} - \mathbf{Y}\|^2, \quad (5)$$

where C is the penalty factor to balance the two terms. Furthermore, we have

$$\boldsymbol{\beta}^* = \begin{cases} \mathbf{H}^\top \left(\mathbf{H}\mathbf{H}^\top + \frac{\mathbf{I}}{C} \right)^{-1} \mathbf{Y}, & n < L \\ \left(\mathbf{H}^\top \mathbf{H} + \frac{\mathbf{I}}{C} \right)^{-1} \mathbf{H}^\top \mathbf{Y}, & n \geq L \end{cases} \quad (6)$$

3.2. Maximum mean discrepancy (MMD)

Maximum Mean Discrepancy (MMD) (Gretton et al., 2006) is

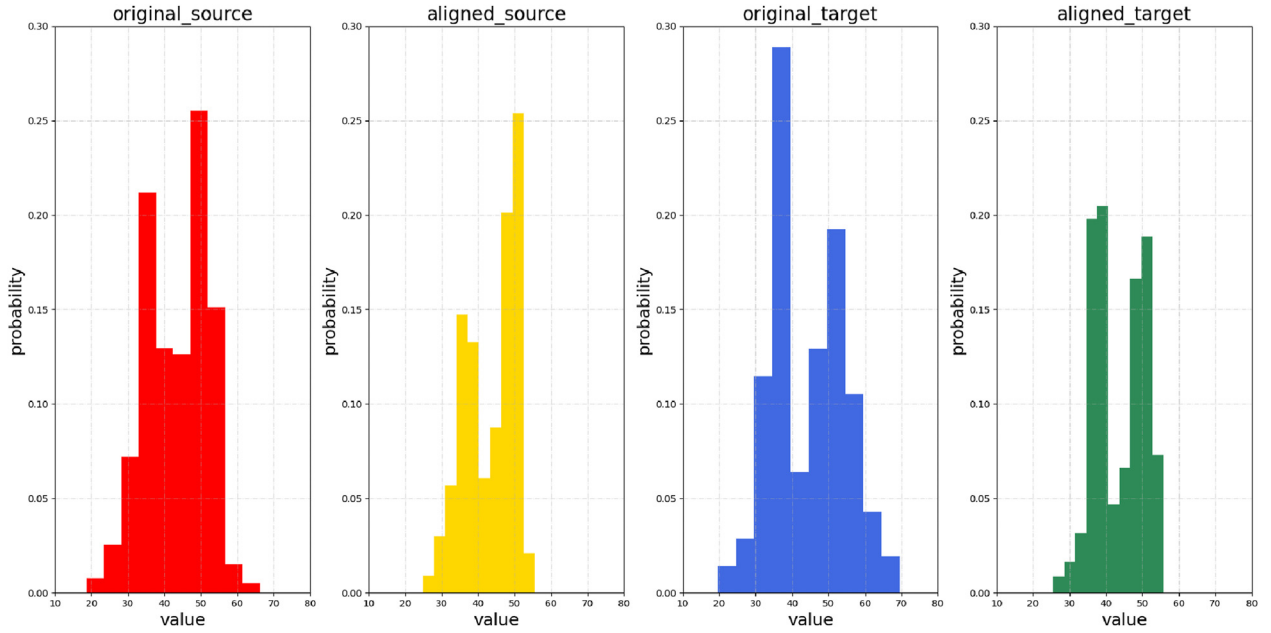


Fig. 9. Frequency distribution histogram of COND log in the task A → B.

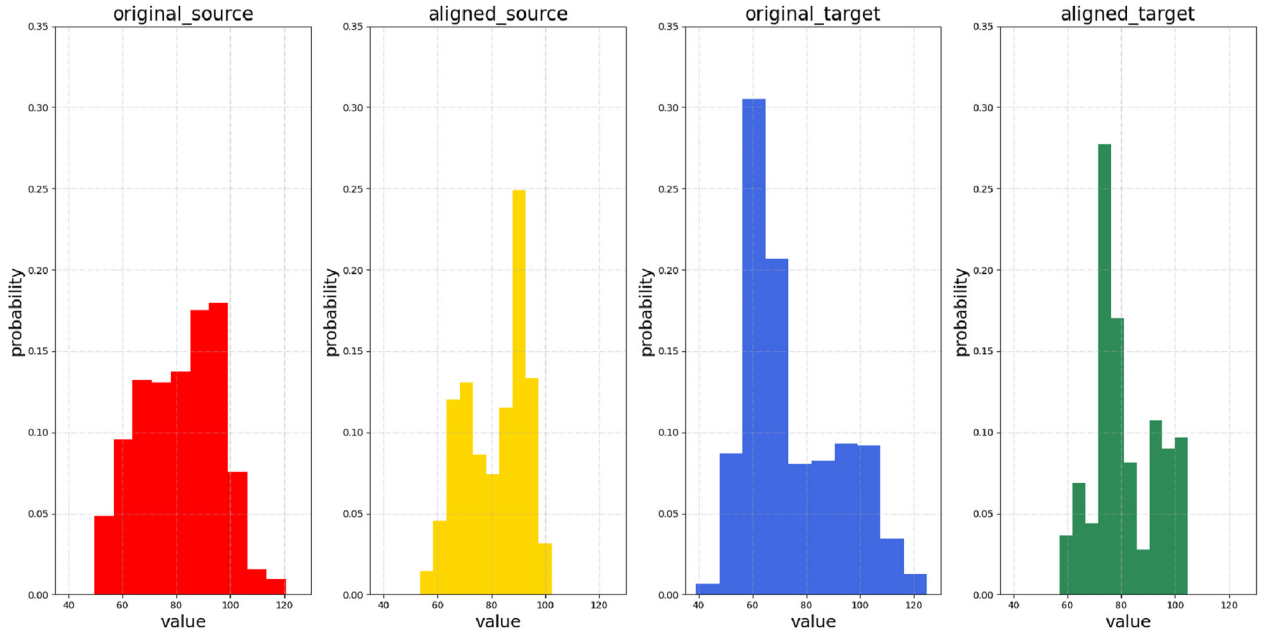


Fig. 10. Frequency distribution histogram of GR log in the task A → B.

generally used to measure the similarity between two probability distributions of the source-domain and the target-domain instances. The distance is measured by mapping the instances into a Reproducing Kernel Hilbert Space (RKHS). Given two distributions P and Q , the formula of the distance based on the maximum mean discrepancy criterion between them is

$$\text{MMD}^2(\mathcal{F}, P, Q) = \sup_{f \in \mathcal{F}} \|E_P[f(x)] - E_Q[f(y)]\|^2 \quad (7)$$

where \mathcal{H} denotes an RKHS space, \mathcal{F} denotes a class of functions. The continuous function $f(\cdot)$ is the mapping function, $\sup(\cdot)$ is an upper bound function, and $E_P(x)$, $E_Q(x)$ are expectations. The empirical estimation of MMD in an RKHS is

$$\text{MMD}^2(X, Y) = \left\| \frac{1}{n} \sum_{i=1}^n \varphi(x_i) - \frac{1}{m} \sum_{j=1}^m \varphi(x_j) \right\|_{\mathcal{H}}^2 \quad (8)$$

where $\varphi(\cdot)$ is the kernel-induced feature mapping function.

3.3. Unilateral Alignment (UA)

For the case that $P_s(\mathbf{x}^s) \neq P_t(\mathbf{x}^t)$, UA of geophysical logs calibration aims to adjust the logs of one well to align it with another. To achieve this goal, we propose the UA calibration method by combining ELM and MMD.

We introduce the random mapping used in ELM to our method, i.e., the instance \mathbf{x}_i is mapped by $\mathbf{h}(\mathbf{x}_i) \in \mathbb{R}^L$, where L is the dimension of mapped instance. Additionally, we define $\mathbf{H}_s \in \mathbb{R}^{n_s \times L}$ and $\mathbf{H}_t \in \mathbb{R}^{n_t \times L}$ as

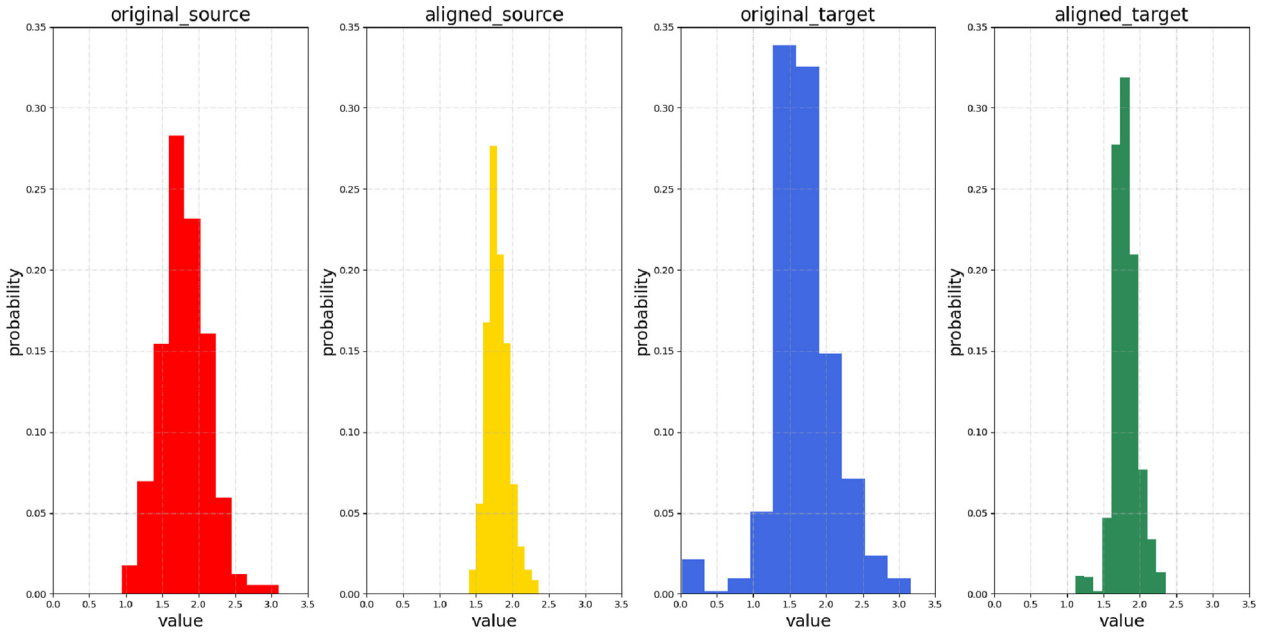


Fig. 11. Frequency distribution histogram of R25 log in the task A → B.

the random mapping matrices of all the instances in datasets D_s , D_t , respectively. In this case, MMD could be expressed by

$$\text{MMD}^2(D_s, D_t) = \left\| \frac{1}{n_s} \sum_{i=1}^{n_s} \mathbf{h}(\mathbf{x}_i^s) - \frac{1}{n_t} \sum_{j=1}^{n_t} \mathbf{h}(\mathbf{x}_j^t) \right\|^2 \quad (9)$$

A transform matrix β operating to each mapped instance is used to reduce the distribution discrepancy, so we have

$$\min_{\beta} \left\| \frac{1}{n_s} \sum_{i=1}^{n_s} \mathbf{h}(\mathbf{x}_i^s) \beta - \frac{1}{n_t} \sum_{j=1}^{n_t} \mathbf{h}(\mathbf{x}_j^t) \beta \right\|^2 = \min_{\beta} \text{Tr}(\beta^T \mathbf{H}^T \Omega \mathbf{H} \beta) \quad (10)$$

where $\text{Tr}(\cdot)$ is the matrix trace, and $\Omega \in \mathbb{R}^{n \times n}$ denotes the marginal PMMD matrix, that is,

$$\tilde{\Omega}_{ij} = \begin{cases} \frac{1}{n_s^2}, & i, j \leq n_s \\ \frac{1}{n_t^2}, & i, j > n_s \\ -\frac{1}{n_s n_t}, & \text{others} \end{cases} \quad (11)$$

Nevertheless, the random mapping function leads to the interaction of logs, thereby the mapped features lose the physical meanings. In this case, the regularization term of source domain maintenance (Chen et al., 2018) is introduced in the process of calculating β , which could maintain the physical meanings of aligned features. Hence, the process of logs calibration is interpretable through our method. Specifically, by introducing the source domain maintenance term, the equation (10) can be written as

$$\min_{\beta} \frac{\lambda}{2} \text{Tr}(\beta^T \mathbf{H}^T \Omega \mathbf{H} \beta) + \frac{\gamma}{2} \|\mathbf{X}_s - \mathbf{H}_s \beta\|^2 \quad (12)$$

where the second term denotes the regularization term of source domain maintenance. $\mathbf{X}_s \in \mathbb{R}^{n_s \times L}$ is the source-domain instance matrix. The penalty factors λ and γ are used to control the contributions of each term. To prevent the overfitting issue, (12) is augmented with a complexity measure term as follows:

$$\min_{\beta} \frac{1}{2} \|\beta\|^2 + \frac{\lambda}{2} \text{Tr}(\beta^T \mathbf{H}^T \Omega \mathbf{H} \beta) + \frac{\gamma}{2} \|\mathbf{X}_s - \mathbf{H}_s \beta\|^2. \quad (13)$$

Eventually, we obtain the solution of (13), that is,

$$\beta^* = \begin{cases} \mathbf{H}_s^T \left(\frac{\mathbf{I}_n}{\gamma} + \mathbf{H}_s \mathbf{H}_s^T + \frac{\lambda}{\gamma} \Omega \mathbf{H} \mathbf{H}^T \right)^{-1} \mathbf{X}_s, & n < L \\ \left(\frac{\mathbf{I}_L}{\gamma} + \mathbf{H}_s^T \mathbf{H}_s + \frac{\lambda}{\gamma} \mathbf{H}^T \Omega \mathbf{H} \right)^{-1} \mathbf{H}_s^T \mathbf{X}_s, & n \geq L \end{cases} \quad (14)$$

where β^* is the optimal transform matrix, and $\mathbf{I}_n \in \mathbb{R}^{n \times n}$, $\mathbf{I}_L \in \mathbb{R}^{L \times L}$ are both identity matrices. Specifically, the UA calibrating function f could be defined as $f(\mathbf{x}^s) = \mathbf{H}_s \beta = \tilde{\mathbf{X}}_s \approx \mathbf{X}_s$ and $f(\mathbf{x}^t) = \mathbf{H}_t \beta = \tilde{\mathbf{X}}_t$, where $\tilde{\mathbf{X}}_s \in \mathbb{R}^{n_s \times d}$, $\tilde{\mathbf{X}}_t \in \mathbb{R}^{n_t \times d}$ are the calibrated instance matrices corresponding to \mathbf{H}_s , \mathbf{H}_t , respectively. In this case, this method could compel the probability distributions $P_s(\tilde{\mathbf{x}}^s) \approx P_t(\tilde{\mathbf{x}}^t)$, and consequently achieve $P_s(\tilde{\mathbf{x}}^s, \mathbf{y}^s) \approx P_t(\tilde{\mathbf{x}}^t, \mathbf{y}^t)$. Subsequently, we could use the source-domain calibrated dataset ($\tilde{\mathbf{X}}_s, \mathbf{Y}_s$) to train an ELM and test on target-domain calibrated dataset $\tilde{\mathbf{X}}_t$ with high accuracy, where $\mathbf{Y}_s \in \mathbb{R}^{n_s \times k_0}$ is the label matrix of source-domain instances. The UA framework is illustrated in Fig. 2.

4. Experimental analysis

In this section, we conduct a large number of experiments using well-logging data collected from multiple wells in Jiyang depression, Bohai Bay Basin to verify the validity and interpretability of our method. First, we describe the dataset and the experimental settings. Second, we compare the performances of UA, S.O., and T.O., and present the accuracy of each lithology via the bar graphs. Finally, we show the visualized results and discuss the interpretability of our method. Here, S.O., means the model trained on source-domain dataset only, which could be seen as the lower bound of UA. If UA has a lower accuracy than S.O., then negative transfer happens. T.O. means the model trained on target-domain dataset only, which could be seen as the upper bound of UA.

4.1. Experimental settings

The dataset consists of three wells, which is shown in Table 1. Wells A, B, and C are from the same area and have the same types of logs and lithologies, that is, they have the same input and output space. To verify the calibration effect, a three-class ELM is trained to classify three types of lithologies and evaluated by average F1-Score which considers both

the accuracy and recall ratios of the classification model. F1-Score could be expressed by

$$\text{averageF1-Score} = \frac{F_1 + F_2 + \dots + F_c}{c} \quad (15)$$

where c is the number of categories and the F_i is F1-Score of i -th category which can be expressed by

$$F_i = 2 \times \frac{P_i \times R_i}{P_i + R_i} \quad (16)$$

where the P_i represents the i -th Precision and R_i is the Recall, which can be computed by

$$\begin{aligned} P_i &= \frac{TP}{TP + FP} \\ R_i &= \frac{TP}{TP + FN} \end{aligned} \quad (17)$$

where the TP, FP and FN represent true positive (an instance are positive classes and are judged to be positive classes), false positive (a false class is judged to be a positive class) and false negative (a false class is judged to be a negative class), respectively.

Due to the wide range of logs, we use min-max normalization to convert all logging data to those within 0 and 1. The number of hidden neurons are set 800, the weights and bias are generated randomly by the Gaussian distribution with variance equalling 3. The trade-off parameter γ is searched from 10^{-3} to 10^4 , and λ from 10^2 to 10^6 . S.O. and T.O. are realized by weighted ELM which could eliminate the influence of unbalanced classes. All experiments are implemented by Python 3.8 on a desktop computer with 2.90-GHz CPU and 16.00-GB RAM.

4.2. Comparative study

The performances of these methods are evaluated by different combinations of wells on the experimental dataset. For example, we set Well A as the source domain for training and Well B as the target domain for testing (i.e., task $A \rightarrow B$). Therefore, there are a total of six combinations to be compared, i.e., $A \rightarrow B$, $A \rightarrow C$, $B \rightarrow A$, $B \rightarrow C$, $C \rightarrow A$, and $C \rightarrow B$. Table 2 lists the macro-average F1-scores of these methods on 6 tasks. We have the following observations. (i) In all six tasks, the high F1-score was achieved by using T.O. method, while the F1-score was significantly decreased by using S.O. method. This indicates that the data distribution varies from well to well. (ii) The F1-scores of UA performing on all tasks are higher than S.O., which shows the effectiveness of UA in the case of data discrepancy. (iii) The F1-scores of UA is very close to those of the T.O. baselines. Especially, it can reach 97.2% on the task $A \rightarrow B$. (iv) In the six tasks, the F1-scores of UA are improved by about 10% compared with S.O. baseline method, and even improved by 20% on the task $B \rightarrow C$. (v) Although the F1-score of S.O. baseline is relatively high, UA still have an improvement on task $A \rightarrow C$. The F1-scores for each lithology are shown in Fig. 3. It is easy to find that S.O. usually overfits to the mudstone so that its F1-score might be high on some tasks. UA eliminates the distribution discrepancies in the six tasks, so the classification after logs calibration could perform well on each lithology.

4.3. Deeper discussions

Taking experiment $A \rightarrow B$ as an example, Fig. 4 and Fig. 5 show the changes of logs of Well A and Well B. It is obvious that each log in source domain has a slight change as shown in Fig. 4. Most of the macroscopic characteristics are preserved. Some microscopic characteristics are weakened due to low-pass filtering function of the regularization term of source domain maintenance. As shown in Fig. 5, some logs has a great change and the predictions are significantly improved by UA. Observing areas a and b in Fig. 5, we think that UA smoothen R25 by integrating

CAL, so the calibrated R25 can be seen as that under the situation without borehole expansion. Such a smoothness in R25 avoids many misclassifications. In areas of c and d, although the calibrated logs change greatly, it keeps the basic variation trend of the original logs.

The following analysis are based on the statistic information of logs, i.e., the differences of means and distribution histograms between two domains. Taking $A \rightarrow B$ and $A \rightarrow C$ as an examples, it can be seen from Table 3 and Table 4 that the mean of each log in the aligned (calibrated) target domain and source domain are almost the same as those in the original source domain. This indicates that the means are aligned unilaterally. Subsequently, Fig. 6 presents the distribution histograms of Well A SP and Well B SP in two domains. It can be seen that the SP log distribution of target domain is obviously aligned to the one of source domain, while the original SP log distribution of Well A and Well B is quite different. This could be the reason why the F1-score in the task $A \rightarrow B$ could be increased greatly as shown in Table 2. According to Fig. 7, it can be seen that there is a small difference in SP log distribution between Well A and Well C, so S.O. method used to predict Well C can achieve a high F1-score of 92.3%. In this case, UA should align to Well A in the high-dimensional space while preserving the original information of Well C as much as possible. The frequency distribution histograms of other logs are shown in Figs. 8–11 which all support the superiority of UA.

5. Conclusions and further work

In this paper, we formulated the geophysical logs calibration problem and give its statistical explanation, and exhibited an interpretable machine learning method, i.e., UA, which could align the logs from one well to another without losing the physical meanings. Massive experiments in 3 wells and 6 tasks showed the effectiveness and interpretability from multiple views. The involved UA method is an unsupervised feature domain adaptation method, so it does not rely on any labels from cores. Actually, the source-domain labels could be used to pre-train a model and outputs some pseudo-labels for the target domain, such that we can obtain a conditional MMD instead of the marginal MMD used in the paper. With the conditional MMD, the conditional probability distribution of both domain could be aligned, which might have a better calibration effect than the current work.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This work was supported in part by the National Natural Science Foundation of China under Grant 61903353, and in part by the SINOPEC Programmes for Science and Technology Development under Grant PE19008-8.

References

- Bergen, K.J., Johnson, P.A., de Hoop, M.V., et al., 2019. Machine learning for data-driven discovery in solid earth geoscience. *Science* 363 (6433).
- Birnie, C., Ravasi, M., Liu, S., Alkhalifah, T., 2021. The potential of self-supervised networks for random noise suppression in seismic data. *Artif. Intell. Geosci.* 2, 47–59.
- Chang, J., Li, J., Kang, Y., Lv, W., Xu, T., Li, Z., Xing Zheng, W., Han, H., Liu, H., 2021a. Unsupervised domain adaptation using maximum mean discrepancy optimization for lithology identification. *Geophysics* 86 (2), ID19–ID30.
- Chang, J., Kang, Y., Zheng, W.X., Cao, Y., Li, Z., Lv, W., Wang, X.-M., 2021b. Active domain adaptation with application to intelligent logging lithology identification. *IEEE Trans. Cybern.*
- Chang, J., Kang, Y., Li, Z., Zheng, W.X., Lv, W., Feng, D.-Y., 2022. Cross-domain lithology identification using active learning and source reweighting. *IEEE Geoscience and Remote Sensing Letters* 19, 1–5.
- Chen, Y., Song, S., Li, S., Yang, L., Wu, C., 2018. Domain space transfer extreme learning machine for domain adaptation. *IEEE Trans. Cybern.* 49 (5), 1909–1922.

- Dunham, M.W., Malcolm, A., Kim Welford, J., 2020. Improved well-log classification using semisupervised label propagation and self-training, with comparisons to popular supervised algorithms. *Geophysics* 85 (1), O1–O15.
- Fouedjio, F., 2020. Exact conditioning of regression random forest for spatial prediction. *Artif. Intell. Geosci.* 1, 11–23.
- Gretton, A., Borgwardt, K., Rasch, M., Schölkopf, B., Smola, A., 2006. A kernel method for the two-sample-problem. *Adv. Neural Inf. Process. Syst.* 19, 513–520.
- Huang, G.-B., Zhu, Q.-Y., Siew, C.-K., 2004. Extreme learning machine: a new learning scheme of feedforward neural networks. In: *Proceedings of the IEEE International Joint Conference on Neural Networks*, vol. 2. Ieee, pp. 985–990.
- Huang, G.-B., Zhou, H., Ding, X., Zhang, R., 2011. Extreme learning machine for regression and multiclass classification. *IEEE Trans. Syst. Man Cybern. B (Cybernetics)* 42 (2), 513–529.
- Imamverdiyev, Y., Sukhostat, L., 2019. Lithological facies classification using deep convolutional neural network. *J. Petrol. Sci. Eng.* 174, 216–228.
- Li, S., Song, S., Huang, G., Wu, C., 2018. Cross-domain extreme learning machines for domain adaptation. *IEEE Trans. Syst. Man Cybern.: Systems* 49 (6), 1194–1207.
- Li, Z., Kang, Y., Feng, D., Wang, X.-M., Lv, W., Chang, J., Zheng, W.X., 2020. Semi-supervised learning for lithology identification using Laplacian support vector machine. *J. Petrol. Sci. Eng.* 195, 107510.
- Li, Z., Kang, Y., Lv, W., Zheng, W.X., Wang, X.-M., 2021. Interpretable semisupervised classification method under multiple smoothness assumptions with application to lithology identification. *Geosci. Rem. Sens. Lett. IEEE* 18 (3), 386–390.
- Liu, H., Wu, Y., Cao, Y., Lv, W., Han, H., Li, Z., Chang, J., 2020. Well logging based lithology identification model establishment under data drift: a transfer learning method. *Sensors* 20 (13), 3643.
- Long, W., Chai, D., Aminzadeh, F., 2016. Pseudo density log generation using artificial neural network. In: *SPE Western Regional Meeting*. OnePetro.
- Lv, W., Kang, Y., Zhao, Y., 2018. Self-tuning asynchronous filter for linear Gaussian system and applications. *IEEE/CAA J. Automat. Sin.* 5 (6), 1054–1061.
- Magrini, F., Jozinović, D., Cammarano, F., Michelini, A., Boschi, L., 2020. Local earthquakes detection: a benchmark dataset of 3-component seismograms built on a global scale. *Artif. Intell. Geosci.* 1, 1–10.
- Ruckebusch, G., 1983. A kalman filtering approach to natural gamma ray spectroscopy in well logging. *IEEE Trans. Automat. Control* 28 (3), 372–380.
- Stojanov, P., Li, Z., Gong, M., Ca, R., Carbonell, J., Zhang, K., 2021. Domain adaptation with invariant representation learning: what transformations to learn? *Adv. Neural Inf. Process. Syst.* 34.
- van der Maaten, L., Hinton, G., 2008. Visualizing data using t-SNE. *J. Mach. Learn. Res.* 9 (Nov), 2579–2605.
- Vapnik, V., 1999. *The Nature of Statistical Learning Theory*. Springer science & business media.
- Wu, Y., Yang, Y., Lv, W., Chang, J., Li, Z., Feng, D., Xu, T., Li, J., 2022. Robust unilateral alignment for subsurface lithofacies classification. *IEEE Trans. Geosci. Rem. Sens.* 60, 1–13.
- Xu, T., Chang, J., Kang, Y., Lv, W., Li, J., Han, H., Liu, H., 2022. Intelligent cross-well sandstone prediction based on convolutional neural network. *IEEE Geoscience and Remote Sensing Letters* 19, 1–5.
- Yang, Q., Zhang, Y., Dai, W., Pan, S.J., 2020. *Transfer Learning*. Cambridge University Press.
- Zhong, R., Johnson, R.L., Chen, Z., 2020. Using machine learning methods to identify coal pay zones from drilling and logging-while-drilling (lwd) data. *SPE J.* 25, 1241–1258, 03.
- Zhou, R., Cai, Y., Zong, J., Yao, X., Yu, F., Hu, G., 2020. Automatic fault instance segmentation based on mask propagation neural network. *Artif. Intell. Geosci.* 1, 31–35.
- Zhu, L., Li, H., Yang, Z., Li, C., Ao, Y., 2018. Intelligent logging lithological interpretation with convolution neural networks. *Petrophysics* 59, 799–810, 06.