



Federated transfer learning for auxiliary classifier generative adversarial networks: framework and industrial application

Wei Guo¹ · Yijin Wang¹ · Xin Chen¹ · Pingyu Jiang¹

Received: 12 July 2022 / Accepted: 1 April 2023 / Published online: 5 May 2023
© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2023

Abstract

Machine learning with considering data privacy-preservation and personalized models has received attentions, especially in the manufacturing field. The data often exist in the form of isolated islands and cannot be shared because of data privacy in real industrial scenarios. It is difficult to gather the data to train a personalized model without compromising data privacy. To address this issue, we proposed a Federated Transfer Learning framework based on Auxiliary Classifier Generative Adversarial Networks named ACGAN-FTL. In the framework, Federated Learning (FL) trains a global model on decentralized datasets of the clients with data privacy-preservation and Transfer Learning (TL) transfers the knowledge from the global model to a personalized model with a relatively small data volume. ACGAN acts as a data bridge to connect FL and TL by generating similar probability distribution data of clients since the client datasets in FL cannot be directly used in TL for data privacy-preservation. A real industrial scenario of pre-baked carbon anode quality prediction is applied to verify the performance of the proposed framework. The results show that ACGAN-FTL can not only obtain acceptable performance on 0.81 accuracy, 0.86 precision, 0.74 recall, and 0.79 F1 but also ensure data privacy-preservation in the whole learning process. Compared to the baseline method without FL and TL, the former metrics have increased by 13%, 11%, 16%, and 15% respectively. The experiments verify that the performance of the proposed ACGAN-FTL framework fulfills the requirements of industrial scenarios.

Keywords Federated transfer learning · Auxiliary classifier generative adversarial networks · Data privacy · Personalized model

Introduction

Presently machine learning is developing rapidly and playing a valuable role in various industrial scenarios and applications (Ray, 2019). This is a result of its excellent capability to extract the implicit function from data. Traditional machine learning trained a centralized or global model with abundant

data, however, it may run into bottlenecks as in industrial scenarios there are many restrictions. First of all, the training data are not easy to obtain and the data volume is often not enough to train a performance machine learning model (Jing et al., 2019). Moreover, there are many policies and regulations for protecting the privacy and security of industrial data. Therefore, the industrial data exist in the form of isolated islands and cannot be shared to train a global model (Liang et al., 2020). Thirdly, a personalized model is needed since the same types of equipment but different production processes and materials may lead to different patterns and implicit functions in industrial scenarios (Kashyap, 2017). How to quickly and accurately train a personalized model is the core target of industrial machine learning.

Federated Transfer Learning (FTL) can be applied to surmount the aforementioned bottlenecks which can satisfy the demands on lack of data, data privacy-preservation, and personalized model. FTL is a kind of framework that organically

✉ Pingyu Jiang
pjjiang@mail.xjtu.edu.cn

Wei Guo
weiguol128@xjtu.edu.cn

Yijin Wang
wangyijin1702@gmail.com

Xin Chen
chenxin9911@stu.xjtu.edu.cn

¹ State Key Laboratory for Manufacturing Systems Engineering, Xi'an Jiaotong University, Xi'an, China

combines Federated Learning (FL) with Transfer Learning (TL) from the system point. FL trains a model on a large corpus of decentralized data with data privacy-preservation and TL can transfer the knowledge from the trained model to a personalized model with a relatively small data volume. According to this idea, there are three FTL paradigms: (1) FL and TL have an ordinal relation. FL enables the creation of a powerful global model with data privacy-preservation. After that TL allows that global model to share knowledge between tasks from different but related domains (Saha & Ahmad, 2021). (2) TL is a part of FL, that TL is adopted as an approach to establish a personalized model, which can decrease the computation and communication costs in the FL system (Wang et al., 2022; Zhang et al., 2022). (3) FTL allows knowledge to be shared with data privacy-preservation of the clients and enables knowledge to be transferred from the source domain to the target domain in a data federation (Liu et al., 2020a). The three FTL paradigms can be chosen according to the industrial scenario and its targets. In most instances, TL requires the global model and data from FL, however, the data from FL cannot be shared for data privacy-preservation. A data bridge to connect FL and TL is the key technology in FTL.

The major concerns of FTL are data privacy-preservation and personalized model building, which is closer to real industrial scenarios than traditional machine learning. In this paper, we faced an industrial problem of quality prediction in the electrolytic aluminum industry, which is a typical binary classification problem in machine learning. The product called pre-baked carbon anode (PCA) or anode carbon block, is produced in the production lines with petroleum coke and petroleum coke as aggregate and modified pitch as a binder. The aggregate is preheated to a specific temperature and then kneaded by a kneader with the melted pitch. The kneaded materials are transported to a moulder for vibro-compacting to form the shape of PCA. The PCA has been baked in a baking furnace, after that it has a stable geometry. The qualified PCAs will be used as the anode material in the aluminum electrolytic cell. The unqualified PCAs and butt anodes will be crushed and reused in the next PCA production. The brief production process of PCA is shown in Fig. 1.

Since the quality of PCA has direct effects on electrolytic aluminum efficiency, the research on the quality prediction of PCA is very meaningful. In this industrial scenario, training a quality prediction model of PCA mainly faced the two following challenges:

- (1) The carbon plants (with one or more production lines) urgently need a machine learning model for quality prediction, however, they are reluctant to share their production data with others or third parties. Since the production data include some sensitive information,

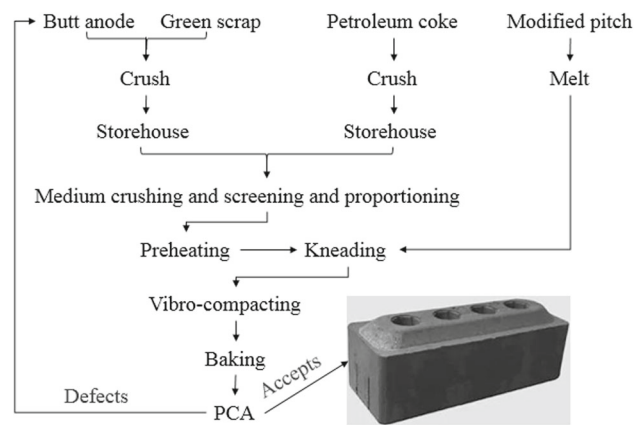


Fig. 1 Production process of the PCA

such as product size, product quantity, production process parameters, etc. There exist practical risks of data leakage and abuse once the data are shared, and the production data of each carbon plant should be protected. A quality prediction model should be trained without data sharing due to privacy and security concerns.

- (2) As the price of primary aluminum has increased, the yield of electrolytic aluminum has increased by 20% from 2016 to 2020 in China. Plenty of production lines have been established. In the real industrial scenario, the production data are not easy to obtain and each newly built production line has its own specific production process. A well-trained quality prediction model may not be appropriate for a new production line. That means a personalized quality prediction model has to be trained with scarce production data. It is hard to train a performance model in that situation based on traditional machine learning methods.

According to the challenges above and the industrial scenario, we focused on the first FLT paradigm which carbon plants (act as the clients) leverage FL to train a global model and then transfer the knowledge of this model to a newly built carbon plant (acts as the target client) by FL. In this paper, a new FTL framework (ACGAN-FTL) based on the idea of Auxiliary Classifier Generative Adversarial Networks (ACGAN) as a data bridge is proposed to tackle the challenges in industrial applications, which considered data privacy-preservation and personalized model.

We summarize our main contributions as follows:

1. A new FTL framework based on ACGAN as a data bridge for industrial applications is proposed, which can train a global model with data privacy-preservation and learn a personalized model ground on the global model through knowledge transfer.

2. We applied the proposed ACGAN-FLT framework to the PCA production with real-life data for quality prediction. Experiments show that ACGAN-FLT noticeably guarantees data privacy and decreases the model training time compared to traditional machine learning approaches which do not apply our framework.
3. The proposed framework is scalable and extensible for personalized model training which focuses on data privacy-preservation. It can not only apply to the PCA quality prediction but also deploy to other industrial applications which consider data privacy-preservation and personalized model-building.

The remainder of this paper is organized as follows. We review the related works and point out some research gaps on product quality prediction, industrial applications of FTL and ACGAN in section “[Related works](#)”. In section “[Proposed ACGAN-FTL framework](#)”, we present the ACGAN-FTL framework and data preprocessing method. In section “[Experiments](#)”, we set up experiments, then analyze and compare the performance of the proposed method with others. The discussions and conclusions are presented in sections “[Discussion](#)” and “[Conclusion](#)”.

Related works

Machine learning for product quality prediction

Product quality improvement is the eternal melody of the industry. Therefore, product quality monitoring (Castaño et al. 2020), prediction (Liu et al., 2021a), or control (Dutta et al., 2021) with AI strategies have been popular and common topics. Their target is identifying defective products and promptly improving product quality with advanced techniques. In this paper, we focused on predicting the quality of PCAs since the inner of a PCA is invisible and each process has an effect on it. Researchers applied various machine learning methods to realize product quality prediction, Neural Network is a common and useful algorithm. Hou et al. (2022) established a prediction model based on back-propagation Neural Network and Genetic Algorithm to predict the quality of the extrusion thread. The effectiveness of the method was validated by an industrial dataset. Chang et al. (2020) proposed a back-propagation Neural Network optimized by Mind Evolutionary Algorithm to predict the penetration quality of asymmetrical fillet root welding. Singh and Desai (2022) developed an image-based framework considering pre-trained Convolutional Neural Network (CNN), ResNet-101 to detect surface defects on the centerless grinding of tapered rollers. Kusuma and Huang (2022) predicted the product quality of silicon steel sheets in terms of the average kerf width of a straight slot cutting by using

the DNN. Since production processes are subject to continuous changes, Tercan et al. (2022) applied memory-aware synapses and weight transfer to realize continual learning of neural networks for quality prediction in injection molding. Wang et al. (2019) proposed a Neural Network model for automatically predicting work-in-progress product quality and applied powder metallurgy process to verify that the model has high accuracy in the quality prediction.

Besides the Neural Network, other machine learning methods are also qualified for quality prediction. Ye et al. (2018) established a weighted Random Forest algorithm to forecast the casting billet quality with considering multi-process parameters. Schorr et al. (2020) presented a Random Forest model based on torque measurements to predict the concentricity as well as the diameter of drilled and reamed bores of the valves. Wang et al. (2020) incorporated Random Forest with Bayesian optimization for tackling the problem of product quality prediction, especially the products have large-scale dimensions data. Liu et al. (2020b) proposed an end-to-end unified product quality prediction framework called QTD and designed a novel bidirectional serial–parallel Long Short-Term Memory (LSTM) algorithm for product quality prediction. Bai et al. (2021) integrated rough knowledge with LSTM to form an ensemble learning framework for manufacturing quality prediction and verified the performance of this framework with a manufacturing quality control dataset. Wang et al. (2021) proposed an LSTM recurrent neural network to accurately predict the quality of ultrasonic welding based on monitored signals. Ren et al. (2021) proposed a semi-supervised parallel deep factorization machine model for product quality prediction. The performance and effectiveness of the proposed methods were demonstrated on a real-world dataset. Jung et al. (2021) tried to test and compare the performances of different machine learning algorithms (tree-based algorithms, regression-based algorithms, and autoencoder) in product quality prediction. Zhang et al. (2021) proposed an enhanced bidirectional graph attention network to predict quality indices of the manufacturing process with considering dependency relationships among multiple machines in the multistage manufacturing process. Ma et al. (2022) proposed a soft sensor modeling framework that applied k-nearest neighbor mutual information for dimension reduction and variable selection and then designed a bidirectional gated recurrent unit structure for quality prediction.

To sum up, the papers discussed above leverage typical machine learning methods to get satisfactory results on product quality prediction by training a model based on a mass of data or samples. However, a mass of data is not easily accessible in most industrial scenarios. Since the data exist in the form of isolated islands and are not easily shared due

Table 1 FTL frameworks and applications

Researchers	Frameworks	Applications	Data bridge	Datasets
Zhang and Li (2021, 2022a, 2022b)	Machinery fault FTL	Fault diagnosis	Prior distribution, Deep Adversarial Networks	CWRU dataset, Crack dataset
Chen et al., (2020, 2021)	FedHealth, FedHealth2	Wearable device	Shared data	Parkinson disease dataset
Li and Wang (2019)	FedMD	Heterogeneous data	Uniquely designed models	FEMNIST, CIFAR10/CIFAR100
Yang et al. (2021)	FedSteg	Image Steganalysis	Shared data	BOSSbase 1.01
Li et al. (2021)	FedSWP	Occupational health and safety	Shared data	YawDD
Majeed et al. (2021)	Cross-Silo FTL	Traffic Classification	Model transfer	UNB ISCX VPN-NonVPN network traffic dataset
Kopparapu and Lin (2021)	TinyFedTL	Tiny device	Shared data	ImageNet
Raza et al. (2022)	ECG-FTL	Electrocardiography	Shared data	Electrocardiography dataset
Cheng et al. (2022)	Intrusion detection FTL	Intrusion detection	Shared data	NSL-KDD1, UNSW-NB15
Fan et al. (2020)	IoT Defender	Intrusion detection	Shared data	CICIDS2017, NSL-KDD
Zhang et al. (2022)	COVID-19 FTL	COVID-19 detection	Shared data	COVID-19 Radiography Database, Chest X-Ray Images

to data privacy-preservation. Moreover, a personalized training scheme should be considered for better product quality prediction.

FTL frameworks and applications in industrial scenarios

FL was first proposed by Google in 2016 and then applied this framework to realize next-word prediction and word completions functions without data leakage (McMahan et al., 2016). FL is a typical decentralized machine learning that clients collaboratively train a global model under the orchestration of a central server, while keeping the training data in clients (Ge et al., 2021). From the aspect of the model, the trained model of FL can be used as a basic model and reused in the same or similar tasks only requiring a few adjustments. However, it is hard to find an available basic model in industrial scenarios. TL is a learning technique to transfer a source domain knowledge into a new domain (target domain) and minimize the mismatch between them (Zhuang et al., 2021). TL can not only improve the learning performance in a new domain but also reduce the model training time and data-labeling cost (Felipe & Anna, 2019). FTL combines the advantages of FL and TL for training a personalized model with data privacy-preservation.

The core idea of FTL is to share the knowledge instead of the data and to transfer the knowledge to other tasks for

building a personalized model. The thoughts of FTL are close to the demands of industrial scenarios, researchers started to focus on the applications of FTL in recent years as Table 1 shown.

In conclusion, researchers have proposed kinds of frameworks for FTL to tackle the problems in relative industrial scenarios. However, there are still two challenges to be resolved.

1. From the data bridge perspective, most researchers assumed that the server has public data or shared data from clients which can be used in TL as source data. However, in most industrial scenarios, the server only acted as an aggregator and did not own data. That means the source data have to be obtained from the clients, which should destroy the privacy-preservation of FL;
2. From the dataset perspective, researchers verified their proposed frameworks by utilizing public datasets which are elaborated. In industrial scenarios, the data usually have useless or irrelative information such as inconsistencies, missing values, noises, or redundancies. In the face of industrial scenarios, public datasets are not the best choice to verify the FTL frameworks when compared with real industrial production data.

In this paper, we used production data which are collected from the equipment embedded with sensors in production

lines and applied the idea of the Generative Adversarial Network to generate similar probability distribution data of the clients' data which can be shared and used as source data in TL.

Auxiliary classifier generative adversarial network

Generative Adversarial Network (GAN) is a machine learning framework that the generator produces the candidate data and the discriminator used for evaluation. They contest with each other in a game until the discriminator cannot evaluate the data generated by the generator (Sun et al., 2021). The contest operates in terms of data distributions, in other words, GAN can generate similar probability distribution data with original data. GAN has been applied in a range of applications, including data augmentation, image synthesis, image superresolution, style transfer, semantic image editing, and classification (Creswell et al., 2018). The datasets in our case study are used for the binary classification task, Auxiliary Classifier Generative Adversarial Network (ACGAN) is applied since it does not only predict if the data is 'real' or 'fake' but also has to provide the 'source' or the 'class label' of the given data (Odena et al., 2017). Besides that, in most industrial scenarios, the proportion of unqualified products is very low, we should intentionally generate these data for tackling the data imbalance problem.

In most industrial scenarios, ACGAN is used to realize data augment and data imbalance. For data augmentation, Zhou et al. (2021) proposed a fully automatic electrocardiogram classification system based on ACGAN. They augmented the data by generating various coupling matrix inputs according to different arrhythmia classes. Li et al. (2022) applied ACGAN as a data-augmentation method to tackle the issue of lacking samples in rotating machinery fault diagnosis. Waheed et al. (2020) applied ACGAN to generate synthetic chest X-ray images for COVID-19 detection since training a model needs a large number of chest X-ray images and collecting images in such a short time is hard. Shao et al. (2019) developed a framework based on ACGAN to learn and generate real one-dimensional raw data from mechanical sensor signals. They used the generated signals as augmented data for machine fault diagnosis. Dixit et al. (2021) proposed a conditional ACGAN framework coupled with model agnostic meta learning to tackle the problem of limited fault samples by generating synthetic samples.

For data imbalance, Liu et al. (2021b) proposed a focal ACGAN for defective wafer pattern recognition with imbalanced data and measured it on a real-world wafer map dataset to prove it outperforms Support Vector Machine and Convolutional Neural Network. Zhao and Yuan (2021) proposed an improved ACGAN with an Auto-Encoder to enhance the fault diagnosis performance with imbalanced data. The Auto-Encoder was used for the similarity estimation of generated

samples. Xia et al. (2018) proposed generating virtual acoustic data using the ACGAN and demonstrated the improved performance of the proposed technique compared to existing acoustic event detection systems. Ding et al. (2022) proposed a tabular data sampling method integrating the K-nearest neighbor method and tabular ACGAN to balance the normal samples and attack samples. Du et al. (2022) applied the variational auto-encoder and ACGAN to generate electrocardiogram images and then the convolutional neural network is used to automatically recognize arrhythmias. Luo et al. (2022) developed a two-stage GAN to address imbalanced data problems in network traffic classification. In the first stage, TimeGAN learned the time-series information of sequence data, and then ACGAN generated synthesized data samples for balancing data in the second stage.

In a summary, the above-mentioned papers have proved that ACGAN is a useful method for data augmentation or data balancing. According to the mechanism of ACGAN, it also could be a feasible approach to be used as a data bridge for data privacy-preservation in FTL. In this paper, we leveraged ACGAN to generate similar probability distribution data instead of original data to train a model in TL. To the best of our knowledge, ACGAN was applied as the data bridge for the first time.

Proposed ACGAN-FTL framework

Problem definition and overview of the framework

The problem of FTL can be defined as: the data from N different clients (carbon plants), denote the clients by $\{C_1, C_2, \dots, C_N\}$ and the target client (the new carbon plant) C_t . The PCA production data they have by $\{D_1, D_2, \dots, D_N\}$ and D_t . The majority of machine learning methods train the target model f_T only by D_t . In this paper, the problem can be defined as gathering all the clients' data $D = \cup_{i=1}^N D_i$ to train a global model f_G where any client C_i does not reveal its data D_i to one another. And the knowledge of the global model can be transferred to build the target model f_T . Let A_{D+D_t} and A_{D_t} represent the accuracy of the target client's model trained by FTL and non-FTL, respectively. And the training time of FTL and non-FTL are T_{D+D_t} and T_{D_t} , it should be pointed out that T_{D+D_t} is only the training time of TL based on trained global model, since if the global model has been trained once, it can be reused many times. The objective of ACGAN-FTL is to guarantee the performance of FTL is superior to that of traditional learning methods, and the training time is far less than traditional learning methods, which can be expressed as:

$$A_{D+D_t} - A_{D_t} > \Delta \quad (1)$$

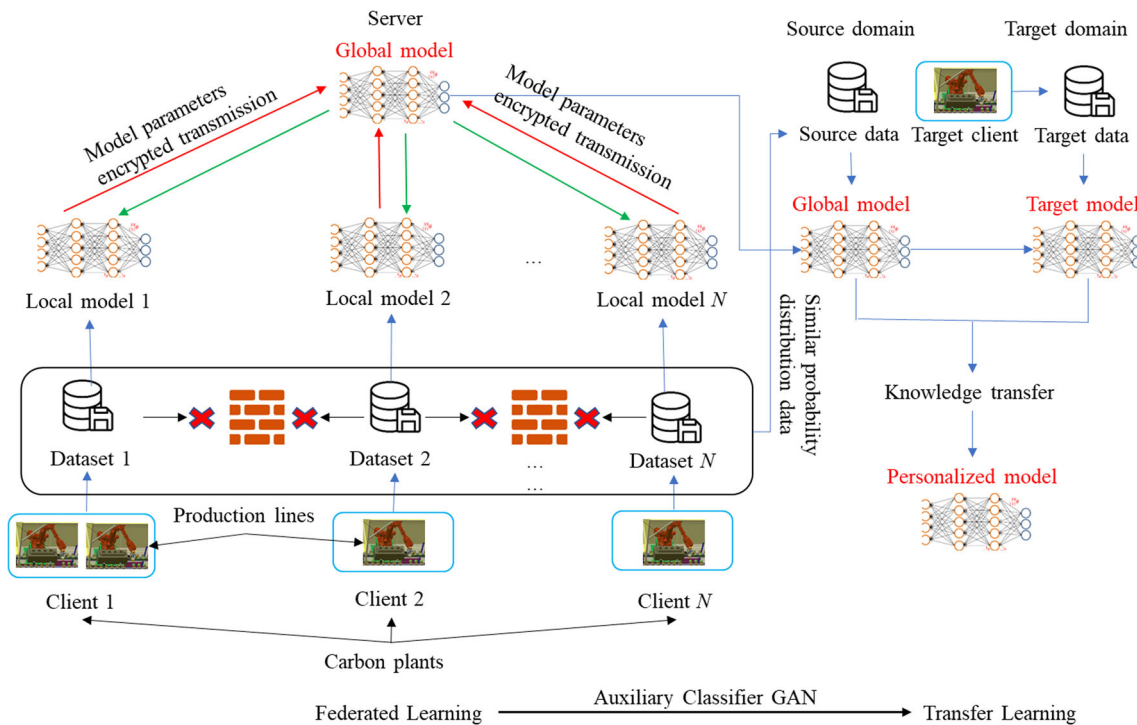


Fig. 2 The framework of ACGAN-FTL

$$T_{D+D_t} - T_{D_t} \ll 0 \tag{2}$$

where Δ is an extremely small nonnegative real number.

According to the industrial scenario and challenges explained above, the FTL framework based on ACGAN for PCA quality prediction is proposed as Fig. 2 shown. The ACGAN-FTL aims at achieving a personalized model with accurate quality prediction for the new carbon plant (target client). There are N clients and one server in FL and a target client t in TL, the data of each client cannot be shared. In the ACGAN-FTL framework, FL and TL have ordinal relation that FL trains a global model and TL transfers the knowledge of global model to the target client for building a personalized model.

In FL, the server sets a Deep Neural Network (DNN) model structure for this case study (the global model in Fig. 3) and then the global model is distributed to all clients where each of them can train their local models with client datasets. Then the encrypted parameters of local models are updated to the server where the parameters (weights and biases) are aggregated and formed the global model. The server distributes the encrypted parameters of the global model to all clients for the next iteration until satisfying preset conditions.

In TL, ACGAN is used as a data bridge to generate similar probability distribution data with client datasets. The generated data replace the client datasets as source data and a personalized model is trained with the global model and target data together. The knowledge of the global

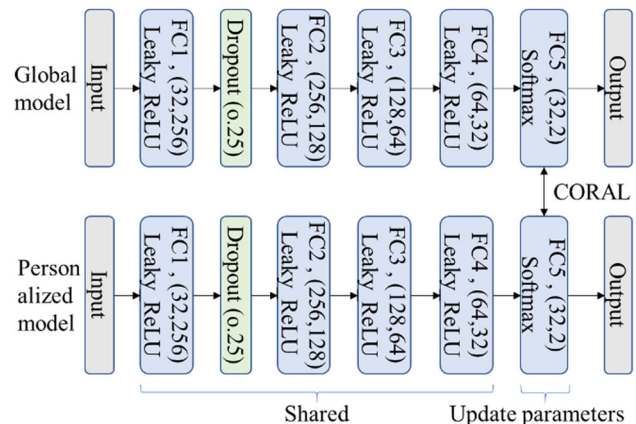
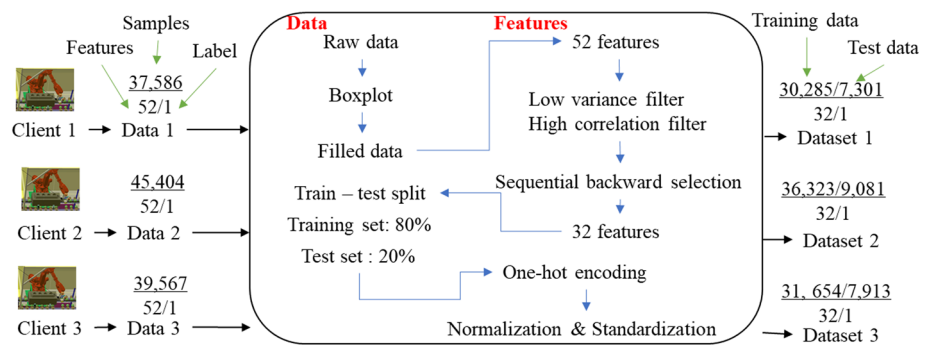


Fig. 3 Knowledge transfer between the global model and personalized model

model is transferred to the target model by appropriate approaches, such as model fine-tune, maximum mean discrepancy, CORAL (correlation alignment), etc. Since the label of target data is unavailable in this case study, the CORAL method is applied to align the domains between the source and target models (Fig. 3).

Data preprocessing and dataset of PCA production

In this paper, we used the data collected from PCA production equipment according to the production process. There are

Fig. 4 Data preprocessing and dataset

three carbon plants (Client1, Client2, Client3) and one server in FL, and one carbon plant (target client) in TL. Client1, Client2 and Client3 have 37,586, 45,404 and 39,567 samples respectively, and each sample includes 52 features collected by sensors on equipment, such as kneaders, moulders, baking furnaces, etc. The target client has about 7800 samples with the same features and no labels. Particularly, there is no data in the server, which is a third-party service provider and only used for the global model aggregation. Beneficial from this setting, all three isolated client datasets can be conducted to evaluate the performance of PCA quality prediction with data privacy-preservation.

In order to obtain knowledge with better performance, data preprocessing plays a significant but also time-consuming role within the entire data mining process (Fan et al., 2021). Industrial raw data usually come with useless or irrelevant information such as inconsistencies, missing values, noises or redundancies. Through conducting proper data exploration and feature engineering, high-quality data are able to be figured out that fit the following mining process. We used the Boxplot to discover data and fill outliers with median values. For the part of feature engineering, a low variance filter and high correlation filter were applied to reduce the low information features and high similar features at begin. After that, a Sequential Backward Selection (SBS) algorithm was used to further reduce features. SBS is categorized as one of the wrapper methods. The entire variable set in SBS is set up as the removal start and the feature will be eliminated at a time if its elimination provides the least impact on the predictor performance (Somol et al., 1999). After implementing previous methods in our case study, the features of PCA production data drop from 52 to 32 as the final input shape. Then the training data and test data were split in a ratio of 80:20. We used one-hot encoding for tackling the nominal data. Finally, normalization and standardization methods were applied. The procedure of dataset establishing is shown in Fig. 4. The same data preprocessing procedure was used to target client data.

ACGAN as a data bridge

ACGAN is proposed by Odena et al. (2017), as an extension of the original GAN for the improved training of GAN for image synthesis. Different from the original GAN, the generator of ACGAN is provided with two parameters on noise Z and label c instead of only noise Z , which attempts to generate the data for that label. The discriminator has to classify whether the data is 'real' or 'fake' and it also has to predict the label of the data. Hence, the objective function (loss function) of ACGAN has two sections: the loglikelihood of the correct source L_S , and the log-likelihood of the correct class L_C .

$$L_S = E[\log P(S = real|X_{real})] + E[\log P(S = fake|X_{fake})] \quad (3)$$

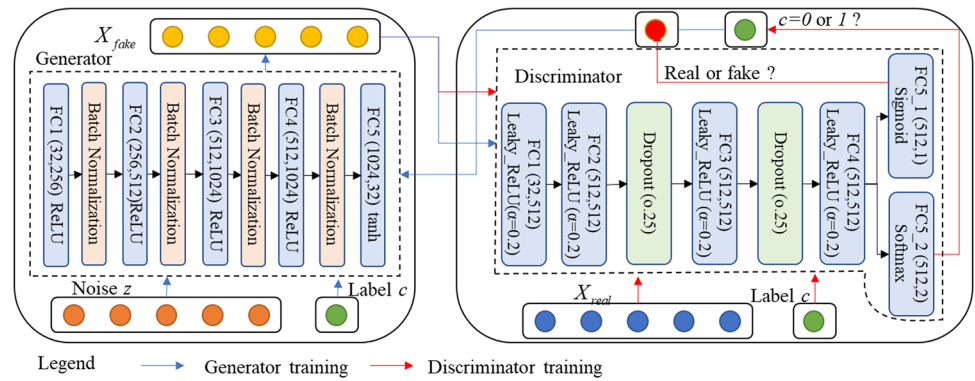
$$L_C = E[\log P(C = c|X_{real})] + E[\log P(C = c|X_{fake})] \quad (4)$$

The Discriminator is trying to maximize its reward ($L_C + L_S$) and the Generator is trying to minimize the Discriminator's reward ($L_C - L_S$). The learning procedure of ACGAN used in this case study is shown in Fig. 5.

ACGAN-FTL learning process

The learning process of ACGAN-FTL is presented in Algorithm 1. This algorithm includes two parts, part A for FL procedure and part B for TL procedure, some key technologies should be explained. In part A, the server established the DNN global model f_G and sent it to each client, and after the local training on each client, its parameters are updated to the server with data privacy-preservation. An additively homomorphic encryption method is selected for parameter encrypted transmission since it allows computation to be implemented directly on encrypted data without decrypting the data, when decrypted, it produces the same output as when the operation was performed on the encrypted data (Hardy

Fig. 5 The learning procedure of ACGAN



et al., 2017). And additively homomorphic encryption has better efficiency when compared with fully homomorphic encryption. After the clients updated model parameters to the server, a model aggregation method called Federated Averaging (FedAvg) was applied to form the global model. The formula of FedAvg is denoted as (Harbnet et al., 2016)

$$w_{t+1} \leftarrow \sum_{n=1}^N \frac{|D_n|}{|D|} w_{t+1}^n \tag{5}$$

where $|D_n|$ and $|D|$ denote the sample size of D_n and D respectively, w_{t+1}^n is the model parameters from the n th client and w_{t+1} is the model parameters after aggregation.

In part B, the knowledge should be transferred from the global model to the personalized model. Each client uses ACGAN to train a generator to generate similar probability distribution data with the client dataset. The task in this case study is classification. Moreover, the labels of the source domain dataset are available and the labels of the target domain dataset are unavailable. The lower layers are frozen and the higher layer (generally refers to the last layer) is updated as Fig. 3 shown since the lower layers of DNN extract generalized features and higher layers extract personalized features. The CORAL method is applied to calculate the distance and minimize the differences between the source and target domains (the output of the last layers of the global model and personalized model), the loss of CORAL can be calculated as follows

$$L_{coral} = \frac{1}{4d^2} \|C_s - C_t\|_F^2 \tag{6}$$

where d is the dimension of embedding features and $\|\cdot\|_F^2$ represents the Frobenius norm. C_s and C_t are the covariance matrices of source and target domains. Besides that, the classification loss of the global model should be calculated since the labels of the target domain dataset are unavailable. The global model f_G by minimizing a loss function is as follows

$$\arg \min_{\Theta} L(\Theta|X) = \frac{1}{n} \sum_{i=1}^n l(c^i, f_G(X_{fake}^i)) \tag{7}$$

where Θ be the parameters of the global model including the weights and biases, $l(\cdot, \cdot)$ represents the classification loss function, $X = \{X_{fake}^i\}_{i=1}^n$ denotes the n synthetic data generated by ACGAN and c^i be the label of X_{fake}^i . The total loss on the target client side can be denoted as

$$L_{total} = \arg \min_{\Theta} L(\Theta|X) + \lambda L_{coral} \tag{8}$$

where $\lambda > 0$ is the tradeoff parameter.

Algorithm 1. The learning process for ACGAN-FTL

Input: Dataset from clients $\{D_1, D_2, \dots, D_N\}$, D_t , and η

Output: Personalized classification model f_T

Part A: FL procedure

- #1: Set an initial global model (DNN) f_G and distribute f_G to all clients
- #2: Train local models and update all local models to the server using additively homomorphic encryption
- #3: Use FedAvg (Eq. 5) to aggregate local models and form the global model f_G and then distribute it to all clients
- #4: Repeat the steps 2 and 3 until satisfying preset conditions

Part B: TL procedure

- #5: Each client uses ACGAN to generate data (Eqs. 3 and 4) and then gather these data into a dataset D_s
- #6: Select data (X_{fake}, c) from the dataset D_s as source data
- #7: Use the source data, target data and global model to train a personalized model (Eq. 8)

Experiments

In this section, detailed experiments are conducted to evaluate and analyze the performance of the ACGAN-FTL on PCA quality prediction. On the one hand, we analyzed the learning efficiency and accuracy of the proposed framework. On the other hand, we compared the performance of the framework with traditional machine learning methods, such as directly

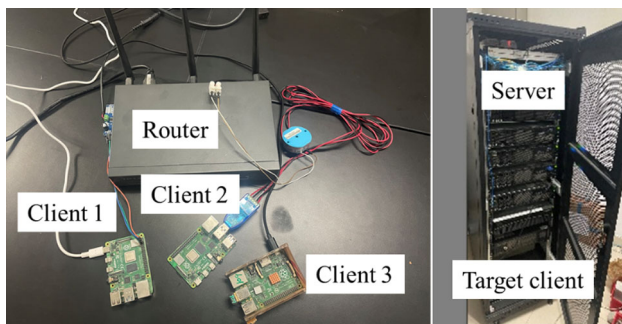


Fig. 6 The hardware configuration of ACGAN-FTL

training a DNN without FTL and transferring a model without ACGAN.

Experiments settings and details

In the experiments, we adopt the production data which collected from four PCA carbon plants (client 1–3 and target client). The data preprocessing and dataset establishing procedure has been explained in section “[Data preprocessing and dataset of PCA production](#)”. For hardware, the ACGAN-FTL is implemented in the three clients and one server for FL, and one target client for TL as Fig. 6 shown. The configuration of the server computer with Inter (R) Xeon (R) CPU E5-2630 at 2.30 GHz, GPU GeForce 1080 Ti and 64 GB RAM. The configurations of clients are raspberry pi 4B with 4 cores arm A72 and 8G RAM. The router is TP-Link WVR450G. For software, we applied Pysyft (0.2.4), Pytorch (1.4.0), Python (3.7) and Raspberry Pi OS (32 bit) based on Debian Buster.

In the procedure of FL, the server firstly distributed the initial DNN to client 1–3 and was utilized for model aggregation by FedAvg. The architecture of the DNN is shown in Fig. 3. On all the clients, the DNN was adopted for training and quality prediction. From all available data obtained from each client, 80% were randomly selected for model training, then the remaining 20% was used for evaluation. We initialized the weights of the first four fully-connected layers in our experiments with “Kaiming normal” instead of random values. As a regularization technique, the dropout was applied once for reducing overfitting by preventing complex co-adaptations on the training data and the probability of an element to be zeroed in the dropout layer was set as 0.25. Some other hyper-parameters were set as follows: the mini-batch size is 128, the learning rate is 0.001 and the Adam method is chosen for the optimization. We adopted additively homomorphic encryption for parameter encrypted transmission.

In the procedure of TL, we chose the ACGAN to generate similar probability distribution data from the three clients respectively in order to remain the data privacy. The architectures of the Generator and Discriminator in our experiment

are shown in Fig. 5. For the optimization, the Adam is chosen with mini-batch size 256, learning rate 0.0001, and beta_1 at 0.5. In addition, batch normalization layers with momentum at 0.8 are applied in the Discriminator to accelerate the training of ACGAN and make it more stable. Especially, the Discriminator obtains two outputs. One represents the sample validity using “Sigmoid” as the activation function and the other means the sample label or class using “Softmax”. For knowledge transfer, the target model was built by the last aggregation with FedAvg on the server. The source domain data were randomly drawn from the samples generated by the ACGAN with a size of 20,000. The last fully-connected layer is changed to be an alignment layer to adopt the CORAL loss for unsupervised domain adaptation. And the remaining layers shared parameters between the source domain and target domain (Fig. 3). They were trained and updated based on the loss combined with classification loss and CORAL loss.

Performance analysis

FL results analysis

FedAvg and additively homomorphic encryption methods have been applied in this FL procedure. We set metrics on accuracy, precision, recall and F1 to evaluate the learning performance (Table 2). The results show that after 25 epochs, the accuracy of each client is basically stable and the average accuracy is 84.14%. The values of precision, recall and F1 achieved the preset target.

ACGAN results analysis

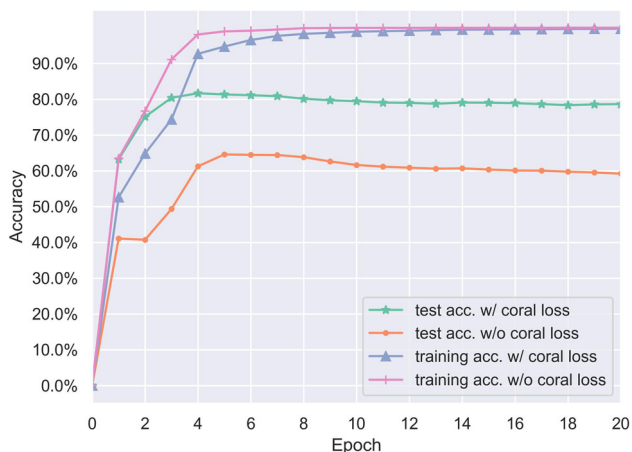
In order to estimate the effectiveness of the proposed ACGAN for data privacy-preservation, a comparison between real data and synthetic data by ACGAN has been taken into account. As shown in Table 3, synthetic data were generated from client datasets on each client respectively. We applied the global model of FL as the testing model. In addition, four evaluation metrics have been chosen for the binary classification problem, namely accuracy, precision, recall and F1. The testing model performs well on the real data since it is trained by client datasets. According to precision and recall scores, each class (qualified or unqualified) obtains balanced prediction results when using real data. In contrast, synthetic data generated by ACGAN shows better prediction ability to detect unqualified PCAs while lower but acceptable power to predict qualified PCAs. Overall, the accuracy score varies by 5% and F1 score differs only by 3%. Thus, synthetic data generated by the proposed ACGAN is able to replace the client datasets for data privacy-preservation in training process.

Table 2 The results of FL on Client1, Client2 and Client3

	Client1	Client2	Client3	Average
Accuracy	0.8380	0.8416	0.8448	0.8414
Precision	0.8492	0.8459	0.8543	0.8498
Recall	0.8240	0.8333	0.8314	0.8295
F1	0.8364	0.8396	0.8427	0.8395

Table 3 A comparison between real data and synthetic data by ACGAN

	Accuracy	Precision	Recall	F1
Real data	0.8345	0.8426	0.8239	0.8331
ACGAN data from client1	0.7776	0.7004	0.9704	0.8136
ACGAN data from client2	0.7710	0.6976	0.9569	0.8069
ACGAN data from client3	0.7794	0.6975	0.9866	0.8173
ACGAN data average	0.7760	0.6985	0.9713	0.8126

**Fig. 7** Training and testing accuracies with v.s. without CORAL loss

TL results analysis

To deeply understand the effectiveness of CORAL method, three plots are generated for elaborating domain shift, where the source domain is about the synthetic data generated by ACGAN. Figure 7 illustrates the training (source domain) and testing (target domain) accuracies for training a personalized model with (w/) v.s. without (w/o) CORAL loss. Beneficial from the consideration of both classification loss and CORAL loss (Eq. 8), not only does the performance on the target domain show better and robust, but the classification accuracy on the source domain also retains a high score (81.1%). In Fig. 8a, two curves represent the trend of the classification loss and the CORAL loss respectively for the training process with consideration of CORAL loss. In Fig. 8b, where the weight λ is set to 0, the CORAL distance between the domains indicates a striking increase for the training process without consideration of CORAL loss.

As a result, it proved that simple fine-tuning without constraints by CORAL loss leads to overfit the features to the source domain and then cannot extract useful features on target domains.

Results comparing

In order to reveal the effectiveness and accuracy of the ACGAN-FTL framework, different methods for the PCA quality prediction problem are implemented for comparisons. Specifically, the following methods are performed in experimental settings similar to the proposed approach.

Baseline

Where not using federated or transfer learning is regarded as the baseline method. This indicates that the target domain model is built only based on its own data without prior knowledge from other clients. It should be pointed out that we assumed the target dataset have labels for the Baseline method.

NoACGAN

As a matter of fact, it remains differences to some degree for model training when utilizing generated data by ACGAN instead of real data for security. In view of the performance of the usefulness of the proposed FTL approach with ACGAN, the NoACGAN method is set up, where the client datasets are directly applied as the source domain. Specifically, the iteratively updated of FL on the clients and server is implemented, and the finally aggregated model on the server is tested on the target client afterward.

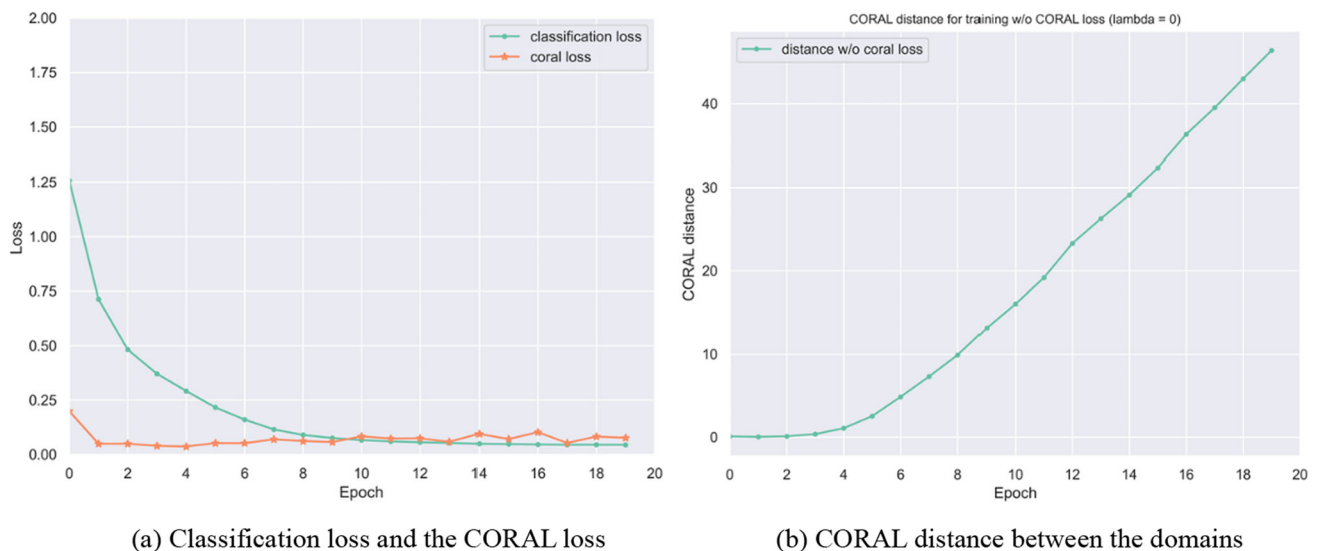


Fig. 8 Domain shift from the source domain to the target domain

Domain adaptation w/o CORAL loss

As a popular TL technique, Deep CORAL considers two aspects of the loss function. One is the classification loss on the source domain to maintain or promote the classification competence on the target domain. The other is the CORAL loss to align both domains. Without considerations of CORAL loss when implementing domain adaptation, the model will be overfitted and perform poorly on the target domain. Thus, the domain adaptation method without CORAL loss is also evaluated in this study. The established model from the proposed FL is fitted to the target client data subsequently with no CORAL loss but only classification loss.

UpperLimit

In order to visualize the gap between the best testing performance and that of the proposed method, the accomplishment of the UpperLimit method is also an essential part of the comparison process. In such a scenario, all client data are generally assumed to be transparent to the server and can be sent with no privacy. Thus, centralized deep learning with data-sharing and domain adaptation with CORAL loss can be implemented by convention. Specifically, the interference of the simulation dataset is excluded so that the domain discrepancy can be clearly minimized when applying CORAL loss with real data.

Figure 9a presents the compared experimental results in terms of testing accuracy for all mentioned methods. About 80% testing accuracies can be attained through the proposed ACGAN-FTL framework in PCA quality prediction. The conventional BaseLine method obtains around 69% testing

accuracies, mainly due to the small amount of data. The NoACGAN method performs slightly better than the proposed ACGAN-FTL method at about 85.1% testing accuracy. The UpperLimit method demonstrates the best prediction results and only is applied for comparison due to the lack of data privacy-preservation, which reaches close to 91% testing accuracy. In short, we can derive the following conclusions: (1) The proposed method is proved to highly exceed the BaseLine method which without conducting federated and transfer learning by 12.1%. It is worth noting that, since there are gaps between source and target domains, such as PCA dimensions, raw materials, and production process parameters, the testing accuracy on the target client is somewhat decreased by 3% than those on FL clients. However, similar patterns are still detected between different domains, that the proposed framework outperforms better the BaseLine method and fine-tuning method without CORAL loss. (2) In spite of 5% higher performance, the NoACGAN approach is still less competitive than the proposed framework since no data privacy-preservation is taken into account, which is not suitable for most industrial scenarios. (3) In short, the comparative results between the proposed framework and the UpperLimit approach are acceptable enough in consideration of several constraints in the industrial scenario. (4) Moreover, the curve of “w/o Coral loss” indicates that without domain adaptation techniques performs distinctly worse than the other approaches on a new but similar PCA production line. Consequently, the effectiveness and superiority of the proposed ACGAN-FTL framework are validated. On the one hand, models with robust and accurate performance can be built owing to the diversity of data samples on different clients in the federated process; On the other hand, overfitting

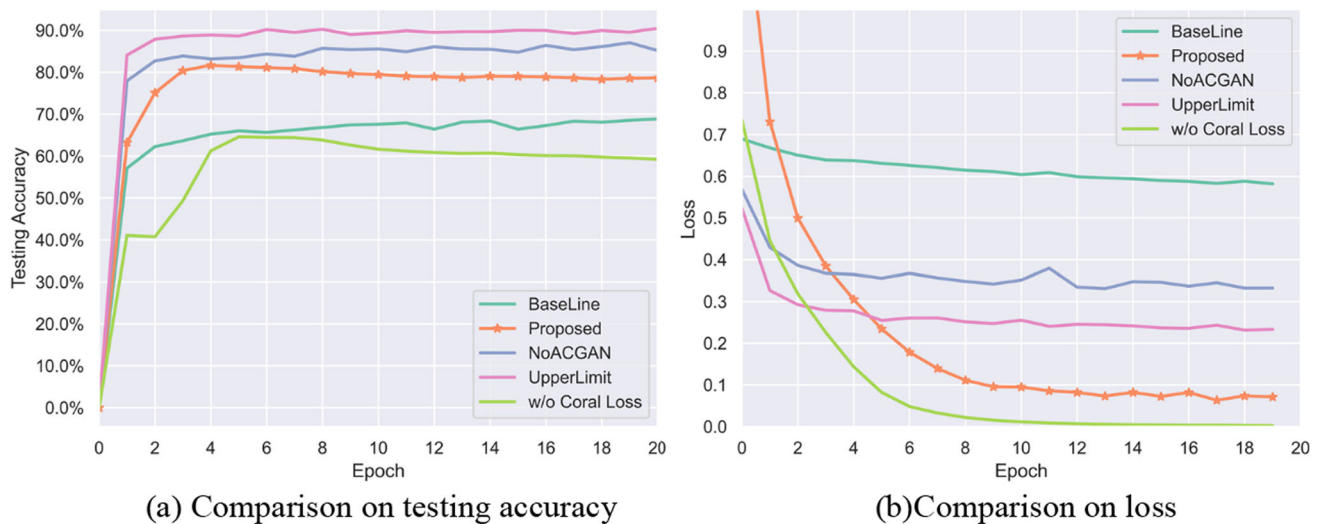


Fig. 9 Comparison of ACGAN-FTL with other methods

has the potential to be avoided when the target data volume is small or owns few labels.

Besides comparison on testing accuracy, Fig. 9b visualizes the loss comparison for all previous methods. Several conclusions can be drawn as follows: (1) It's time-consuming and requires more prior knowledge when training a model from scratch in the BaseLine method. It remains difficulty to gain desired results in a same period of time. (2) The proposed ACGAN-FTL framework can achieve significant improvement due to the combination of classification loss and CORAL loss. Even though the loss of "w/o Coral loss" is lower and decreases more rapidly than the proposed method, it indicates that the global model overfits the features on the source domain rather than the target domain.

Figure 10 illustrates the confusion metrics with both absolute values and percentages to validate the effectiveness of our approach for binary classification in industrial scenarios. Because the price of PCA is relatively high in this production line, it is desired that the false positive be as low as possible to avoid discarding too many qualified PCA resulting in high production costs and low productivity. In our proposed method, the false positives in Fig. 10 up to 9% are acceptable, which means 9% of qualified PCA have been recognized as parts with the defect. Meanwhile, the false negatives reach 23% due to the conflict of precision-recall (Fig. 10). Since in this production line, additional manual inspection is usually added in the last step, the missed unqualified PCA can be further detected. After doing preliminary prediction with proposed artificial intelligence, it can effectively reduce the false negative rate when supplemented with manual inspection for further accurate judgment. Therefore, our proposed method could be deemed as a tool for human-centered augmented intelligence to help decision-making rather than fully automated prediction.

To understand more intuitively the information contained in the confusion metrics, we provide a detailed analysis of ACGAN-FTL via various classification metrics under the previously compared solutions. The precision, recall, and macro F1 score are all computed to give a thorough view of the results. As illustrated in Fig. 11, we can clearly see that the ACGAN-FTL method can not only achieve better accuracy, but also reach better precision, recall, and F1 scores than the BaseLine method. The gaps between the proposed framework and UpperLimit as well as NoACGAN are small and acceptable. The UpperLimit approach only represents an ideal scenario in manufacturing. All client data are assumed to be sent with no privacy and shared on the server. In such a scenario, the global server can simply train a centralized neural network with a large volume of data rather than scarce specialized data in each client. The proposed method and the NoACGAN approach have to obey data privacy in FL so that only model parameters can be sent and gathered instead of real data. Thus, the model performance from UpperLimit surpasses both previous methods. Without the deviation from synthetic data and transport privacy, the domain discrepancy can be clearly minimized using real data as the source domain in Coral loss. In short, few factories have the ability to share data so that the comparative results between the proposed framework and the UpperLimit approach are acceptable enough in consideration of several constraints in the industrial scenario.

Since qualified PCA is the most common class in quality prediction, the precision score is slightly higher than the recall score of our proposed framework. We speculate the reasons as followings: ACGAN learns better and extracts more features for the qualified PCA class than the unqualified PCA class. At the same time, the unqualified PCA class has a little bit fewer samples and more complicated for learning. All in

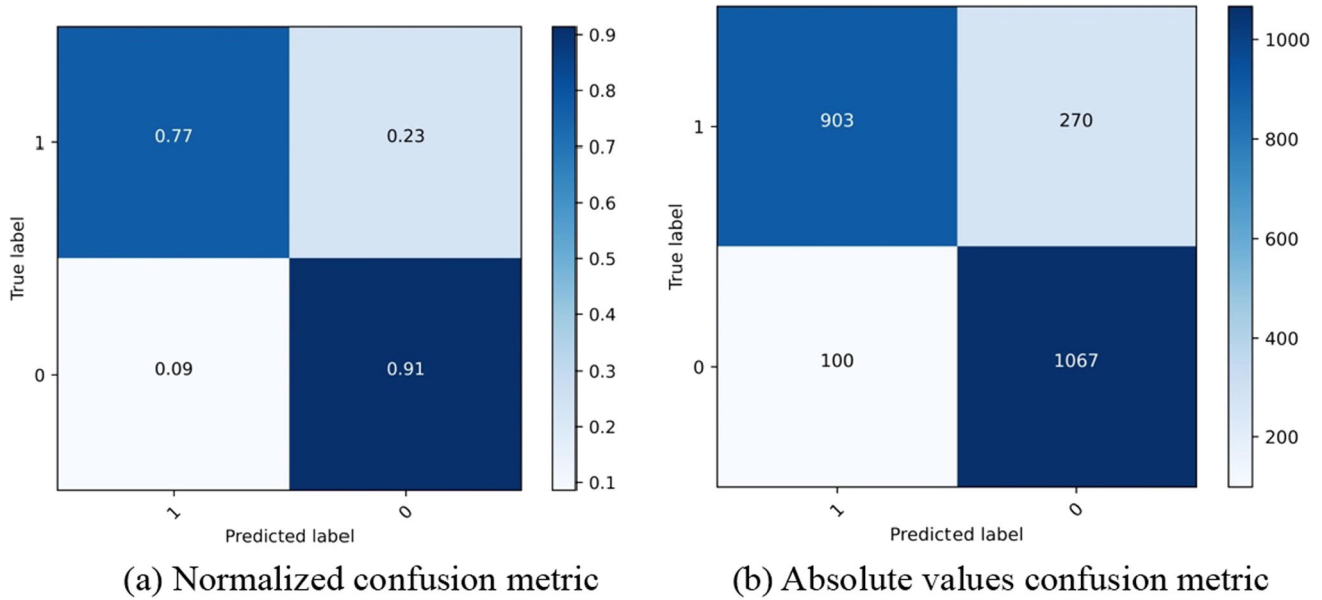


Fig. 10 Confusion metric with both absolute values and percentages

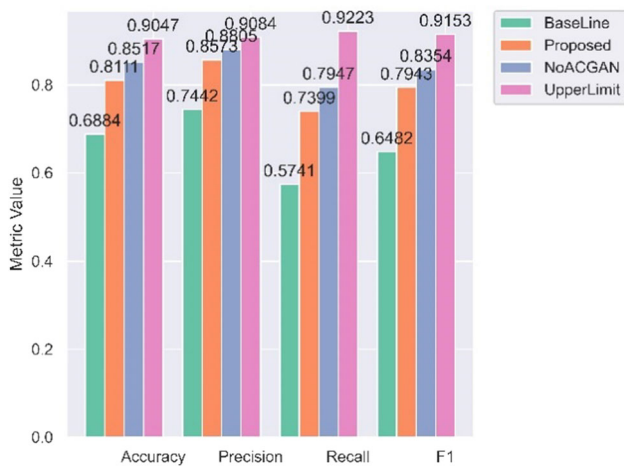


Fig. 11 Comparison of ACGAN-FTL with other methods on accuracy, precision, recall and F1

all, the ACGAN-FTL framework still obtains accurate ability of quality prediction in spite of constrains of data privacy-preservation meanwhile preventing training a model from scratch when facing a new similar PCA production line.

Discussion

ACGAN-FTL is a framework for preserving data privacy and building personalized production models of different clients for quality prediction in the field of industry. Compared with previous findings, four aspects of the proposed ACGAN-FTL’s novelty can be mainly summarized as follows:

Firstly, the ACGAN-FTL framework provides a privacy-preservation solution with double insurance of FL and ACGAN to stimulate different production lines or plants to cooperate for promoting productivity even with scarce data. In previous works in other fields like occupational health and safety, some researchers prefer to assume the server in FL has shared data from clients which can be used in TL as source data. On the contrary, it’s necessary and crucial to comply with data privacy and avoid data sharing in the field of manufacturing so that our proposed approach meets the mentioned needs of the industry and advances the further implementation of this FTL framework in the production sector through ACGAN for generating synthetic data. In Li et al. (2021), the accuracy of their approach could reach 80–85% according to different internal model structures while the accuracy of our proposed method can also achieve 81% even with synthetic data by ACGAN.

Secondly, the ACGAN-FTL framework outperforms current conventional machine learning when meeting a new relative production line. This FTL framework can obtain a more personalized and accurate model through domain adaptation. Extracting useful information from prior knowledge, our proposed method can prevent a cold start to some degree and doesn’t need to train from scratch for saving time and cost.

Thirdly, our proposed method proved the FTL framework has compatibility and feasibility in a more realistic and complete production environment. In Zhang et al. (2021, 2022), this machinery fault FTL has achieved about 95% accuracy with the CWRU dataset which is relatively clean with limited noise, and about 90% accuracy with the crack dataset with

a test rig that consists of two simple bearings. In contrast, even though the accuracy of our proposed method is 81%, its data acquisition was from a real-life PCA production line with more complexity and interference.

Fourthly, we proposed a framework which integrated FL and TL to build a personalized model. The ACGAN-FTL framework can be likened to building frame structure, various algorithms and methods can be applied to this framework for different industrial scenarios. We developed a DNN model for a binary classification problem in this paper. When faced with multiclass classification problems, the model structure of the client and server should be different from the used herein. For example, training models in FL can apply not only the DNN model but also Random Forest (Liu et al., 2021c), Convolutional Neural Network (Alzubi et al., 2022) and etc. According to the features of data, various GANs can be applied to generate corresponding synthetic data. Then the trained model and synthetic data are used to transfer knowledge. Therefore, the proposed ACGAN-FTL framework can be extended for the multiclass classification problems.

Conclusion

This paper proposed a novel federated transfer learning framework based on ACGAN-FTL for pre-baked carbon anode quality prediction in a new production line. ACGAN-FTL organic combines FL with TL through ACGAN as the data bridge from the system perspective. FL aggregates the production data from different production lines without leaking the privacy of the data and achieves the global model which can be used in TL. For considering data privacy-preservation, the production data cannot be used in TL directly, we applied ACGAN to generate similar probability distribution data (synthetic data). The global model, synthetic data (source domain) and target data are applied to achieve a personalized model by knowledge transfer. Experiments have verified the effectiveness and accuracy of the framework by comparing it with BaseLine, NoACGAN, domain adaptation w/o CORAL loss, and UpperLimit methods from the aspects of testing accuracy and loss.

From what have discussed in the paper, we can find that ACGAN-FTL framework is a suitable approach to training a personalized model with considering data privacy-preservation for industrial scenarios. In summary, this work addresses knowledge learning in a data privacy-preservation and personalized model way that makes it easy to understand and operate. The success of this framework is due to the organic combination FL with TL based on ACGAN.

Despite the above innovations, our study still has several limitations which should be considered in the future. We proposed a scalable and extensible ACGAN-FTL framework, which can be applied to other industrial scenarios in

the future. In this paper, the experiments were conducted with a binary classification problem based on discrete production data. However, the multiclass classification problem and sequential data such as time-series data are common in industrial scenarios. Hence, the ACGAN-FTL should be modified to satisfy the new demands, for example, ACGAN is not suitable for generating time-series data, Time-series Generative Adversarial Networks (TimeGAN) (Yoon et al., 2019) or other GAN can generate sequential data should be applied in the future framework.

Acknowledgements The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This research is supported by the National Key Research and Development Program of China under Grant No. 2021YFE0116300 and the Natural Science Foundation for Young Scientists of Shaanxi Province under Grant No. 2021JQ-037.

References

- Alzubi, J. A., Alzubi, O. A., Singh, A., & Ramachandran, M. (2022). Cloud-IIoT based electronic health record privacy-preserving by CNN and blockchain-enabled federated learning. *IEEE Transactions on Industrial Informatics*, 19(1), 1080–1087.
- Bai, Y., Xie, J., Wang, D., Zhang, W., & Li, C. (2021). A manufacturing quality prediction model based on AdaBoost-LSTM with rough knowledge. *Computers & Industrial Engineering*, 155, 107227. <https://doi.org/10.1016/j.cie.2021.107227>
- Castaño, F., Guerra, R., Mohammed, W. M., Nejman, M., & Lastra, J. (2020). Quality monitoring of complex manufacturing systems on the basis of model driven approach. *Smart Structures and Systems*, 26(4), 495–506.
- Chang, Y., Yue, J., Guo, R., Liu, W., & Li, L. (2020). Penetration quality prediction of asymmetrical fillet root welding based on optimized BP neural network. *Journal of Manufacturing Processes*, 50, 247–254.
- Chen, Y., Lu, W., Wang, J., & Qin, X. (2021). FedHealth 2: Weighted federated transfer learning via batch normalization for personalized healthcare. arXiv preprint. <https://arxiv.org/abs/2106.01009>
- Chen, Y., Qin, X., Wang, J., Yu, C., & Gao, W. (2020). FedHealth: A federated transfer learning framework for wearable healthcare. *IEEE Intelligent Systems*, 35(4), 83–93. <https://doi.org/10.1109/MIS.2020.2988604>
- Cheng, Y., Lu, J., Niyato, D., Lyu, B., Kang, J., & Zhu, S. (2022). Federated transfer learning with client selection for intrusion detection in mobile edge computing. *IEEE Communications Letters*, 26(3), 552–556. <https://doi.org/10.1109/LCOMM.2022.3140273>
- Creswell, A., White, T., Dumoulin, V., Arulkumaran, K., Sengupta, B., & Bharath, A. A. (2018). Generative adversarial networks: An overview. *IEEE Signal Processing Magazine*, 35(1), 53–65. <https://doi.org/10.1109/MSP.2017.2765202>
- Ding, H., Chen, L., Dong, L., Fu, Z., & Cui, X. (2022). Imbalanced data classification: A KNN and generative adversarial networks-based hybrid approach for intrusion detection. *Future Generation Computer Systems*, 131, 240–254.
- Dixit, S., Verma, N. K., & Ghosh, A. K. (2021). Intelligent fault diagnosis of rotary machines: Conditional auxiliary classifier GAN coupled with meta learning using limited data. *IEEE Transactions on Instrumentation and Measurement*, 70, 1–11.
- Du, C., Liu, P. X., & Zheng, M. (2022). Classification of imbalanced electrocardiosignal data using convolutional neural network. *Computer Methods and Programs in Biomedicine*, 214, 106483.

- Dutta, G., Kumar, R., Sindhwani, R., & Singh, R. K. (2021). Digitalization priorities of quality control processes for SMEs: A conceptual study in perspective of Industry 4.0 adoption. *Journal of Intelligent Manufacturing*, 32(6), 1679–1698.
- Fan, C., Chen, M., Wang, X., Wang, J., & Huang, B. (2021). A review on data preprocessing techniques toward efficient and reliable knowledge discovery from building operational data. *Frontiers in Energy Research*, 9(2021), 1–17. <https://doi.org/10.3389/fenrg.2021.652801>
- Fan, Y., Li, Y., Zhan, M., Cui, H., & Zhang, Y. (2020). IoTDefender: A federated transfer learning intrusion detection framework for 5G IoT. In *2020 IEEE 14th international conference on big data science and engineering (BigDataSE)* (pp. 88–95). <https://doi.org/10.1109/BigDataSE50710.2020.00020>
- Felipe, L. D. S., & Anna, H. R. C. (2019). A survey on transfer learning for multiagent reinforcement learning systems. *Journal of Artificial Intelligence Research*, 64, 645–703.
- Ge, N., Li, G. H., Zhang, L., & Liu, Y. (2021). Failure prediction in production line based on federated learning: An empirical study. *Journal of Intelligent Manufacturing*. <https://doi.org/10.1007/s10845-021-01775-2>
- Harbnet, M., Eider, M., Daniel, R., & Blaise, A. Y. A. (2016). Federated learning of deep networks using model averaging. arXiv preprint. <https://arxiv.org/abs/1602.05629>
- Hardy, S., Henecka, W., Ivey-Law, H., Nock, R., Patrini, G., & Smith, G., & Thorne B. (2017). Private federated learning on vertically partitioned data via entity resolution and additively homomorphic encryption. arXiv preprint. <https://arxiv.org/abs/1711.10677>
- Hou, H., Zhang, G., Chen, X., & Zhao, Y. (2022). Quality prediction of internal thread cold extrusion based on genetic algorithm optimized BP neural network. *Advances in Mechanical Engineering*, 14(3), 1–10. <https://doi.org/10.1177/16878132221089148>
- Jing, Q., Wang, W., Zhang, J., Tian, H., & Chen, K. (2019). Quantifying the performance of federated transfer learning (pp. 1–7). arXiv preprint. <https://arxiv.org/abs/1912.12795>
- Jung, H., Jeon, J., Choi, D., & Park, J. (2021). Application of machine learning techniques in injection molding quality prediction: Implications on sustainable manufacturing industry. *Sustainability (Basel, Switzerland)*, 13(8), 4120. <https://doi.org/10.3390/su13084120>
- Kashyap, P. (2017). Industrial applications of machine learning. In *Machine learning for decision makers: cognitive computing fundamentals for better decision making* (pp. 189–233).
- Kopparapu, K., & Lin, E. (2021). TinyFedTL: Federated transfer learning on tiny devices. arXiv preprint. <https://arxiv.org/abs/2110.01107>
- Kusuma, A., & Huang, Y. (2022). Product quality prediction in pulsed laser cutting of silicon steel sheet using vibration signals and deep neural network. *Journal of Intelligent Manufacturing*. <https://doi.org/10.1007/s10845-021-01881-1>
- Li, D., & Wang, J. (2019). FedMD: Heterogenous federated learning via model distillation. arXiv preprint. <https://arxiv.org/abs/1910.03581>
- Li, W., Zhong, X., Shao, H., Cai, B., & Yang, X. (2022). Multi-mode data augmentation and fault diagnosis of rotating machinery using modified ACGAN designed with new framework. *Advanced Engineering Informatics*, 52, 101552.
- Li, X., Chi, H., Lu, W., Xue, F., Jianhuan, Z., & Li, C. Z. (2021). Federated transfer learning enabled smart work packaging for preserving personal image information of construction worker. *Automation in Construction*, 128, 1–12.
- Liang, Y., Guo, Y., Gong, Y., Luo, C., Zhan, J., & Huang, Y. (2020). An isolated data island benchmark suite for federated learning. arXiv e-prints. <https://arxiv.org/abs/2008.07257>
- Liu, C., Wang, K., Wang, Y., & Yuan, X. (2021a). Learning deep multimanifold structure feature representation for quality prediction with an industrial application. *IEEE Transactions on Industrial Informatics*, 18(9), 5849–5858.
- Liu, J., Zhang, F., Yang, B., Zhang, F., Gao, Y., & Wang, H. (2021b). Focal auxiliary classifier generative adversarial network for defective wafer pattern recognition with imbalanced data. In *The 5th IEEE electron devices technology & manufacturing conference (EDTM)* (pp. 1–3).
- Liu, Y., Kang, Y., Xing, C., Chen, T., & Yang, Q. (2020a). A secure federated transfer learning framework. *IEEE Intelligent Systems*, 35(4), 70–82. <https://doi.org/10.1109/MIS.2020.2988525>
- Liu, Y., Ma, Z., Yang, Y., Liu, X., Ma, J., & Ren, K. (2021c). Revfir: Enabling cross-domain random forest training with revocable federated learning. *IEEE Transactions on Dependable and Secure Computing*. <https://doi.org/10.1109/TDSC.2021.3104842>
- Liu, Z., Zhang, D., Jia, W., Lin, X., & Liu, H. (2020b). An adversarial bidirectional serial-parallel LSTM-based QTD framework for product quality prediction. *Journal of Intelligent Manufacturing*, 31(6), 1511–1529.
- Luo, W., Yang, W., He, J., Huang, H., Chi, H., Wu, J., & Shen, Y. (2022). Fault diagnosis method based on two stage GAN for data imbalance. *IEEE Sensors Journal*, 22(22), 21961–21973.
- Ma, L., Wang, M., & Peng, K. (2022). A novel bidirectional gated recurrent unit-based soft sensor modeling framework for quality prediction in manufacturing processes. *IEEE Sensors Journal*, 22(19), 18610–18619.
- Majeed, U., Hassan, S. S., & Hong, C. S. (2021). Cross-silo model-based secure federated transfer learning for flow-based traffic classification. In *The international conference on information networking (ICOIN)* (pp. 588–593).
- McMahan, H. B., Moore, E., Ramage, D., Hampson, S., & Arcas, B. A. Y. (2016). Communication-efficient learning of deep networks from decentralized data. In *The artificial intelligence and statistics, PMLR* (pp. 1273–1282).
- Odena, A., Olah, C., & Shlens, J. (2017). Conditional image synthesis with auxiliary classifier GANs. In *The international conference on machine learning, PMLR* (pp. 2642–2651).
- Ray, S. (2019). A quick review of machine learning algorithms. In *The international conference on machine learning, big data, cloud and parallel computing (COMITCon)* (pp. 35–39).
- Raza, A., Tran, K. P., Koehl, L., & Li, S. (2022). Designing ECG monitoring healthcare system with federated transfer learning and explainable AI. *Knowledge-Based Systems*, 236, 1–17.
- Ren, L., Meng, Z., Wang, X., Zhang, L., & Yang, L. T. (2021). A data-driven approach of product quality prediction for complex production systems. *IEEE Transactions on Industrial Informatics*, 17(9), 6457–6465. <https://doi.org/10.1109/TII.2020.3001054>
- Saha, S., & Ahmad, T. (2021). Federated transfer learning: Concept and applications. *Intelligenza Artificiale*, 15(1), 35–44.
- Schorr, S., Möller, M., Heib, J. O. R., & Bähre, D. (2020). Quality prediction of drilled and reamed bores based on torque measurements and the machine learning method of random forest. *Procedia Manufacturing*, 48, 894–901.
- Shao, S., Wang, P., & Yan, R. (2019). Generative adversarial networks for data augmentation in machine fault diagnosis. *Computers in Industry*, 106, 85–93.
- Singh, S. A., & Desai, K. A. (2022). Automated surface defect detection framework using machine vision and convolutional neural networks. *Journal of Intelligent Manufacturing*. <https://doi.org/10.1007/s10845-021-01878-w>
- Somol, P., Pudil, P., Novovicova, J., & Paclik, P. (1999). Adaptive floating search methods in feature selection. *Pattern Recognition Letters*, 20(11), 1157–1163. [https://doi.org/10.1016/S0167-8655\(99\)00083-5](https://doi.org/10.1016/S0167-8655(99)00083-5)

- Sun, S., Hu, X., & Liu, Y. (2021). An imbalanced data learning method for tool breakage detection based on generative adversarial networks. *Journal of Intelligent Manufacturing*. <https://doi.org/10.1007/s10845-021-01806-y>
- Tercan, H., Deibert, P., & Meisen, T. (2022). Continual learning of neural networks for quality prediction in production using memory aware synapses and weight transfer. *Journal of Intelligent Manufacturing*, 33(1), 283–292.
- Waheed, A., Goyal, M., Gupta, D., Khanna, A., Al-Turjman, F., & Pinheiro, P. A. C. R. (2020). CovidGAN: Data augmentation using auxiliary classifier GAN for improved covid-19 detection. *IEEE Access*, 8, 91916–91923.
- Wang, B., Li, Y., Luo, Y., Li, X., & Freiheit, T. (2021). Early event detection in a deep-learning driven quality prediction model for ultrasonic welding. *Journal of Manufacturing Systems*, 60, 325–336.
- Wang, G., Ledwoch, A., Hasani, R. M., Grosu, R., & Brintrup, A. (2019). A generative neural network model for the quality prediction of work in progress products. *Applied Soft Computing*, 85, 105683.
- Wang, K., Zhou, X., Liang, W., Yan, Z., & She, J. (2022). Federated transfer learning based cross-domain prediction for smart manufacturing. *IEEE Transactions on Industrial Informatics*, 18(6), 4088–4096. <https://doi.org/10.1109/TII.2021.3088057>
- Wang, T., Wang, X., Ma, R., Li, X., Hu, X., Chan, F. T. S., & Ruan, J. (2020). Random forest-Bayesian optimization for product quality prediction with large-scale dimensions in process industrial cyber-physical systems. *IEEE Internet of Things Journal*, 7(9), 8641–8653. <https://doi.org/10.1109/JIOT.2020.2992811>
- Xia, X., Togneri, R., Sohel, F., & Huang, D. (2018). Auxiliary classifier generative adversarial network with soft labels in imbalanced acoustic event detection. *IEEE Transactions on Multimedia*, 21(6), 1359–1371.
- Yang, H., He, H., Zhang, W., & Cao, X. (2021). FedSteg: A federated transfer learning framework for secure image steganalysis. *IEEE Transactions on Network Science and Engineering*, 8(2), 1084–1094. <https://doi.org/10.1109/TNSE.2020.2996612>
- Ye, X., Wu, X., & Guo, Y. (2018). Real-time quality prediction of casting billet based on random forest algorithm. In *2018 IEEE International Conference on Progress in Informatics and Computing (PIC)* (pp. 140–143).
- Yoon, J., Jarrett, D., & Van der Schaar, M. (2019). Time-series generative adversarial networks. In *Advances in neural information processing systems*, 32 (pp. 1–11).
- Zhang, A. S., & Li, N. F. (2022a). When accuracy meets privacy: Two-stage federated transfer learning framework in classification of medical images on limited data: A COVID-19 case study (pp. 1–11). arXiv Preprint. <https://arxiv.org/abs/2203.12803>
- Zhang, D., Liu, Z., Jia, W., Liu, H., & Tan, J. (2021). Path enhanced bidirectional graph attention network for quality prediction in multistage manufacturing process. *IEEE Transactions on Industrial Informatics*, 18(2), 1018–1027.
- Zhang, W., & Li, X. (2021). Data privacy preserving federated transfer learning in machinery fault diagnostics using prior distributions. *Structural Health Monitoring*. <https://doi.org/10.1177/14759217211029201>
- Zhang, W., & Li, X. (2022b). Federated transfer learning for intelligent fault diagnostics using deep adversarial networks with data privacy. *IEEE/ASME Transactions on Mechatronics*, 27(1), 430–439. <https://doi.org/10.1109/TMECH.2021.3065522>
- Zhang, Z., He, N., Li, D., Gao, H., Gao, T., & Zhou, C. (2022). Federated transfer learning for disaster classification in social computing networks. *Journal of Safety Science and Resilience*, 3(1), 15–23.
- Zhao, B., & Yuan, Q. (2021). Improved generative adversarial network for vibration-based fault diagnosis with imbalanced data. *Measurement*, 169, 108522.
- Zhou, Z., Zhai, X., & Tin, C. (2021). Fully automatic electrocardiogram classification system based on generative adversarial network with auxiliary classifier. *Expert Systems with Applications*, 174, 1–13. <https://doi.org/10.1016/j.eswa.2021.114809>
- Zhuang, F., Qi, Z., Duan, K., Xi, D., Zhu, Y., Zhu, H., et al. (2021). A comprehensive survey on transfer learning. *Proceedings of the IEEE*, 109(1), 43–76. <https://doi.org/10.1109/JPROC.2020.3004555>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.