

Review

Deep Unsupervised Domain Adaptation with Time Series Sensor Data: A Survey

Yongjie Shi , Xianghua Ying *  and Jinfa Yang 

School of Artificial Intelligence, Peking University, No. 5 Yiheyuan Road, Haidian District, Beijing 100871, China; shiyongjie@pku.edu.cn (Y.S.); jinfayang@pku.edu.cn (J.Y.)

* Correspondence: xhying@pku.edu.cn

Abstract: Sensors are devices that output signals for sensing physical phenomena and are widely used in all aspects of our social production activities. The continuous recording of physical parameters allows effective analysis of the operational status of the monitored system and prediction of unknown risks. Thanks to the development of deep learning, the ability to analyze temporal signals collected by sensors has been greatly improved. However, models trained in the source domain do not perform well in the target domain due to the presence of domain gaps. In recent years, many researchers have used deep unsupervised domain adaptation techniques to address the domain gap between signals collected by sensors in different scenarios, *i.e.*, using labeled data in the source domain and unlabeled data in the target domain to improve the performance of models in the target domain. This survey first summarizes the background of recent research on unsupervised domain adaptation with time series sensor data, the types of sensors used, the domain gap between the source and target domains, and commonly used datasets. Then, the paper classifies and compares different unsupervised domain adaptation methods according to the way of adaptation and summarizes different adaptation settings based on the number of source and target domains. Finally, this survey discusses the challenges of the current research and provides an outlook on future work. This survey systematically reviews and summarizes recent research on unsupervised domain adaptation for time series sensor data to provide the reader with a systematic understanding of the field.

Keywords: deep learning; unsupervised domain adaptation; time series sensor data; survey



Citation: Shi, Y.; Ying, X.; Yang, J. Deep Unsupervised Domain Adaptation with Time Series Sensor Data: A Survey. *Sensors* **2022**, *22*, 5507. <https://doi.org/10.3390/s22155507>

Academic Editor: Quazi K. Hassan, Ebrahim Ghaderpour and Spiros Pagiatakis

Received: 26 June 2022

Accepted: 21 July 2022

Published: 23 July 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Sensors capture physical parameters from the observed environment and convert them into observable electrical pulses [1–4]. A wide variety of sensors are used in a wide range of applications in manufacturing and machinery [5–7], transportation [8–13], healthcare [14–18] and many other aspects of our daily lives. In the field of mechanical engineering, for example, accelerometer sensors placed around gearboxes or bearings can capture the vibration signals of a machine and predict possible impending failures [7]. In the healthcare field, voltage signals from a patient's brain are captured by placing voltage sensors on the patient's brain and used to identify patient commands [16], etc. By continuously recording physical parameters over a period of time, the current operating state of the monitored system can be analyzed and unknown risks can be assessed.

In recent years, the analysis techniques for time series data have been rapidly developed. Commonly used models for processing time series data are spectral and wavelet analysis [19–21], 1D Convolutional Neural Networks (CNNs) [22–24], Recurrent Neural Networks (RNNs) represented by Long Short-Term Memory (LSTM) networks [25–27] and Gated Recurrent Unit (GRU) [28,29], and the recently emerged Transformer [30]. Deep models have a large number of parameters and powerful feature extraction capabilities, and the parameters of the models are optimized and updated by loss functions and a gradient back propagation [31,32]. However, deep models are often based on the assumption that

the distribution of training and test data is similar. In practice, however, this requirement is not satisfied. For example, in the fault diagnosis of rolling bearings, the vibration signals collected by accelerometer sensors under different loads (e.g., rotation speed) have different patterns. If vibration signals collected under one operating condition are used to train a deep network and data collected under another load condition is used to test the trained model, the performance of the model may be significantly degraded. As another example, in the brain-computer interface, a model trained with a large number of signals collected on one person is tested on another person and the model does not perform well. This is because the performance of the model tends to degrade when the model is trained on one scene (source domain) and tested on data collected on another scene (target domain). One way to solve this problem is to label a large amount of labeled data in a new scene and retrain the deep model. However, relabeling the data for each new scene requires significant human resources. Due to the ability to easily obtain unlabeled data in the target domain, many researchers have adopted unsupervised domain adaptation (UDA) techniques to improve the performance of the model in the target domain. That is, the model is trained using labeled data in the source domain and a large amount of unlabeled data in the target domain.

This survey collects papers on unsupervised domain adaptation with time series sensor data from academic studies in 2018 to date. First, this survey divides them by industry, listing the background of different applications, sensors often used, the sources of data discrepancies, and the commonly used research datasets. Then, this survey classifies these different methods for unsupervised domain adaptation for time series data, including domain adaptation methods aligned in the input space, feature space, and output space, and model-based domain adaptation methods. After that, the paper classifies the existing domain adaptation settings according to the number of sources and target domains and analyzes their advantages. Finally, the survey summarizes the existing methods and looks into possible future research directions. Unlike previous reviews [33–35] that focus on a particular research direction, this paper takes sensors that collect time series data as a starting point and surveys different industries that need to monitor the status of their systems and summarizes the sensors they use and the reasons why domain gaps exist. A classification of deep learning approaches to address these domain gaps is presented, as well as the advantages of different adaptation settings. The aim is to provide the reader with a comprehensive and systematic understanding of deep unsupervised domain adaptation for time series sensor data.

The remainder of the paper is structured as follows (see Figure 1). Section 2 explains some basic concepts. Section 3 divides the collected papers by industry. Section 4 divides the studies according to different adaptation methods. Section 5 divides the existing methods by domain adaptation settings. Section 6 discusses the present study and provides an outlook. Section 7 makes a conclusion.

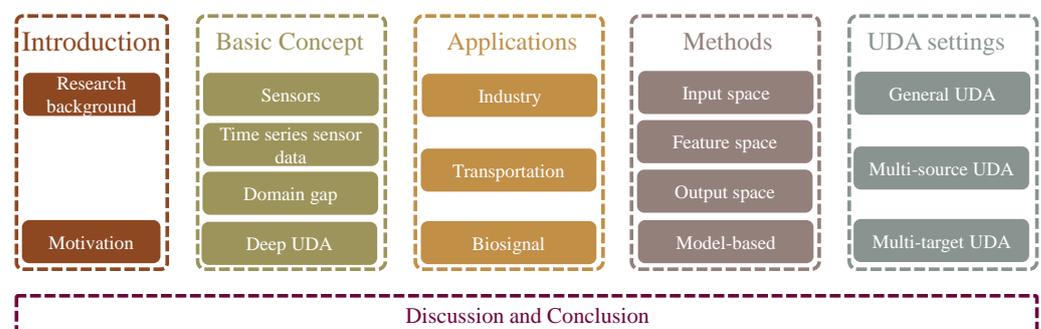


Figure 1. The organization of this survey.

2. Basic Concept

To facilitate the understanding of this paper, this section explains some basic concepts, including what a sensor is, the definition of time series sensor data, what a domain is, and what domain adaptation and deep unsupervised domain adaptation are.

2.1. Sensors

A sensor converts physical phenomena into a measurable digital signal, which can then be displayed, read, or processed further [1]. According to the physical characteristics of sensor sensing, common sensors include temperature sensors that sense temperature, acceleration sensors that sense motion information, infrared sensors that sense infrared information, etc. According to the way of sensing signals, sensors can be divided into active sensors and passive sensors. Active sensors need an external excitation signal or a power signal. On the other hand, passive sensors do not require any external power and produce an output response. LiDAR is an example of an active sensor, as it requires an external light source to emit a laser. By receiving the returned beam, the time delay between emission and return is calculated to determine the distance to an object. Passive sensors, such as temperature sensors, acceleration sensors, and infrared sensors, do not require external excitation and can directly measure the physical characteristics of the system being monitored. A wide variety of sensors are used in different industries and have greatly increased the productivity of society.

2.2. Time Series Sensor Data

The time series of length n observed by the sensor can be expressed as

$$\mathbf{x} = [x_1, x_2, \dots, x_t, \dots, x_n], \quad (1)$$

where the data point x_t is the data observed by the sensor at moment t . When \mathbf{x} is univariate time series data, x_t is a real value and $x_t \in \mathbb{R}$. When \mathbf{x} is multivariate time series data, x_t is a vector and $x_t \in \mathbb{R}^d$, where the d indicates the dimension of x_t . Much of the time series data collected in practical applications are multivariate data that may be obtained from multiple attributes of a sensor or multiple sensors. For example, in the fault diagnosis of rolling bearings, an acceleration time series of three axes XYZ can be obtained simultaneously by a single accelerometer. Another example is in fault diagnosis of power plant thermal system, where multidimensional time series data are obtained simultaneously by using multiple sensors, such as temperature sensors, pressure sensors, flow rate sensors, etc.

2.3. Domain Gap

Data collected under a certain distribution is referred to as a domain, and this distribution can be understood as a specific data collection scenario [36]. For example, in human behavior recognition, inertial sensor data collected at the arm is referred to as one domain, while inertial sensor data collected at the leg is referred to as another domain. The data collected in different scenarios differ in distribution, and this difference is known as the domain gap (see Figures 2 and 3).

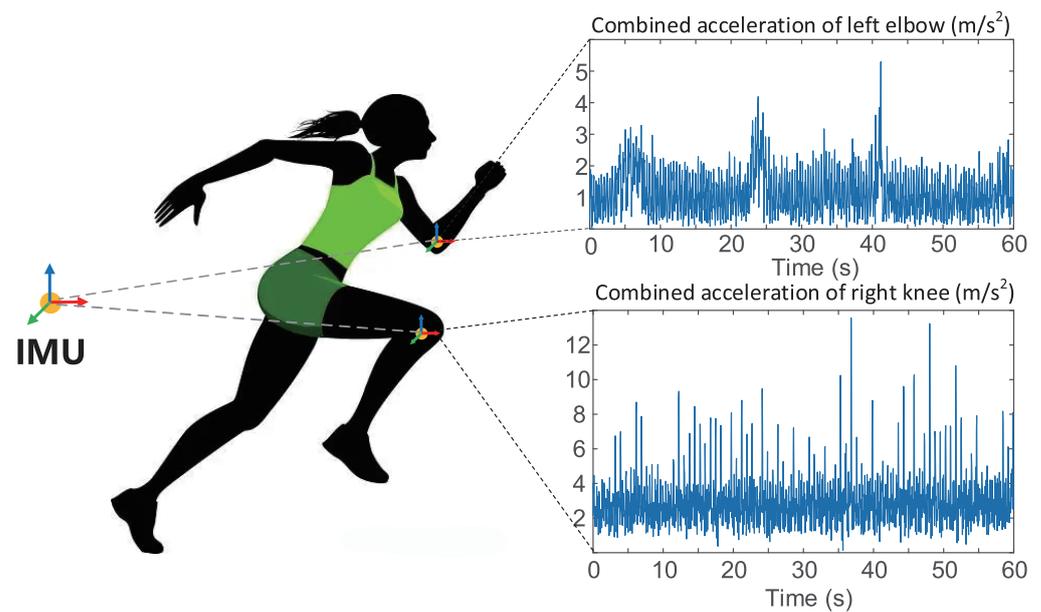


Figure 2. Diagram of human activity recognition based on inertial measurement unit (IMU) [37]. Two IMUs are placed on the arm and leg respectively to record the motion pattern at each position for further recognition of human motion. As can be seen from the figure, there are some differences in the data collected from different limbs. Reprinted with permission from Ref. [37]. Copyright 2020 IEEE.

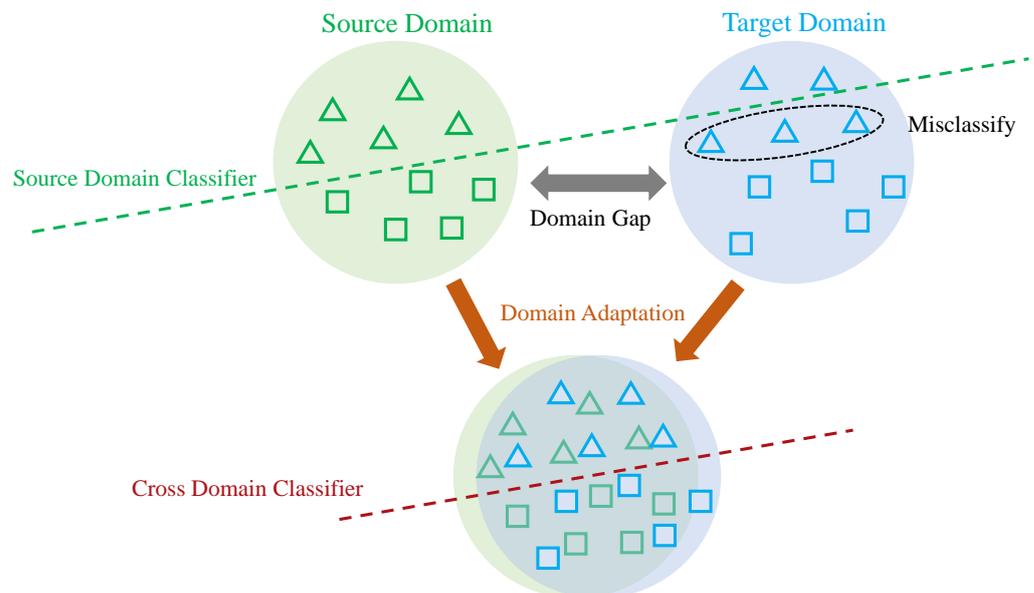


Figure 3. Illustration of source and target data with original feature distributions (**top**), and new feature distributions (**bottom**) after domain adaptation, where domain adaptation techniques help to alleviate the “domain shift” problem between source and target domains.

2.4. Deep Unsupervised Domain Adaptation

In the concept of domain adaptation, scholars usually refer to the training set as the source domain and the test set as the target domain. Assuming that the model is trained with data collected at the arm (source domain) and later tested with data collected at the leg (target domain), the model does not perform as well. This is because the deep learning model can only recognize test data with the same distribution as the training data. One way to improve the performance of the model in the target domain is to label a large amount of data in the target domain and fine-tune the model trained in the source domain using

a supervised approach. However, labeling a large amount of data requires significant human resources.

Domain adaptation (see Figure 3) refers to adapting the model trained in the source domain to the target domain, i.e., reducing the domain gap between the source and target domains and improving the performance of the target domain. Domain adaptation includes semi-supervised domain adaptation and unsupervised domain adaptation. Semi-supervised domain adaptation requires that samples from the target domain are partially labeled. In contrast, unsupervised domain adaptation does not require the target samples to have labels. Deep unsupervised domain adaptation specifically refers to improving the adaptation ability of deep learning models in the target domain. Due to the complexity and nonlinearity of deep learning models, a large number of algorithms for deep unsupervised domain adaptation have emerged in recent years, including mapping-based algorithms, adversarial learning-based algorithms, etc.

3. Applications

Although there are many papers dedicated to UDA of time series sensor data, they have different application backgrounds and research directions. This section systematically analyzes the applications of UDA with time series sensor data in recent years and broadly divides them into three application areas, namely industry, transportation, and biosignal. Each field contains several specific application directions. Moreover, this section summarizes in detail which sensors are included in each application direction, what causes the domain gaps, and which datasets are publicly available. The overall classification results are shown in Table 1, and this section will present them in detail in different subsections.

Table 1. Classification results based on different applications.

	Applications	Sensors	Domain Gap	Datasets	References
Industry	fault diagnosis of rolling bearings	accelerometer, microphone	different working conditions, different positions of the sensors, different machines	CWRU [38], PU [39], IMS [40]	[41–59]
	fault diagnosis of power plant thermal system	temperature sensors, pressure sensors, flow rate sensors, etc.	different fault severity, different load conditions	-	[60,61]
	diagnosis of ball screw failure	accelerometer	different positions of the sensors	-	[62,63]
Transportation	capacity estimation of lithium-ion batteries	current sensor, voltage sensor, temperature sensor	different charging/discharging protocols, different in cell type and manufacture	NASA Battery [64], CALCE Battery [65]	[66,67]
	remaining useful life estimation	accelerometer, temperature sensor, pressure sensor, flow rate sensor, etc.	different operating conditions, different failure modes	C-MAPPS [68], IEEE PHM Challenge 2012 bearing dataset [69]	[70–81]
	gearbox fault diagnosis	accelerometer	different operating conditions	-	[82–86]

Table 1. Cont.

	Applications	Sensors	Domain Gap	Datasets	References
Biosignal	EEG based brain-computer interface	EEG electrodes	different subjects, different sessions	SEED [87], DEAP [88], BCI Competition IV-IIa [89], BCI Competition IV-IIb [90]	[91–98]
	human activity recognition	accelerometers, gyroscopes, wifi sensor	different body parts, different users, different sensors	RealWorld [99], Opportunity [100], HHAR [101], PAMAP2 [102]	[103–110]
	EMG based muscle-computer interface	EMG electrodes	different sessions, different subjects	CapgMyo [111], NinaPro [112], CSL-HDEMG [113]	[111,114–117]
	gait analysis	accelerometers and gyroscopes, IR sensors, radar, camera, EMG electrodes	different positions of sensors, different subjects, different moving states	Daphnet [118], OU-ISIR [119], CASIA-B [120], CASIA-C [121]	[122–127]

3.1. Industry

This section introduces several applications of UDA in industry, including fault diagnosis of rolling bearings, fault diagnosis of power plant thermal systems, and diagnosis of ball screw faults.

3.1.1. Fault Diagnosis of Rolling Bearings

Background. A rolling bearing is a precision mechanical component that changes the sliding friction between the running shaft and the shaft seat, thereby reducing frictional losses [128–131]. Rolling bearings are a widely used mechanical component in industry and often have extremely demanding performance standards. Machine failures can be catastrophic and lead to expensive downtime. Without effective diagnostics, one cannot make reliable predictions about when a failure will occur. Therefore, effective fault diagnosis is essential in the industry.

Sensors. Common types of sensors used for rolling bearing signal acquisition are accelerometer sensors [48,52] as well as microphones [46]. Accelerometer sensors measure the health of machinery by being fixed directly near the machinery to collect the vibration signals generated by the mechanical vibrations. Microphones, on the other hand, capture the operating conditions of machinery by collecting the sound waves generated by mechanical vibrations. The signal collected by the microphone will contain more noise than the accelerometer sensor.

Domain Gap. The reasons for the inconsistent distribution of training and test data come from three main sources: different working conditions, different locations of sensors, and different machines.

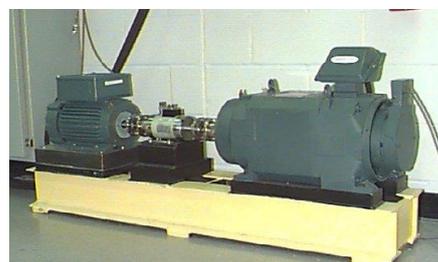
- (1) *Different Working Conditions.* Due to the influence of speed, load, temperature, etc., working conditions often vary during the monitoring period. Collected signals may contain domain shift, which means that the distribution of data may differ significantly under different working conditions. The aim of deep unsupervised domain adaptation-based fault diagnosis is that the model trained using signals under one working condition can be transferred to signals under another different working condition.
- (2) *Different Locations of Sensors.* Since sensors installed on the same machine are often responsible for monitoring different components, sensors located near the fault component are more suitable to indicate the fault information. However, key components have different probabilities of failure rates, leading to the situation where signals from

different locations have different numbers of labeled data. The aim of unsupervised domain adaptation-based fault diagnosis is that the model trained with plenty of labeled data from one location can be transferred to the target domain with unlabeled data from other locations.

- (3) *Different Machines*. For testing cost and safety reasons, it is difficult to collect sufficient samples of marked failures from test machines. In addition, sufficient labeled data can be generated from dynamic simulations or fault seeding experiments. However, the distribution of data from dynamic simulations or fault seeding experiments is different from that of real machines due to similar structure and measurement situations, which is one of the sources of the domain gap.

Datasets. Some common open-source rolling bearing fault diagnosis datasets are the Case Western Reserve University (CWRU) Dataset [38] and the Paderborn University (PU) Dataset [39].

- (1) *CWRU* [38]. The CWRU dataset provided by the Case Western Reserve University Bearing Data Center is one of the most famous open-source datasets in intelligent fault diagnosis. CWRU contains one normal bearing and three fault types, including inner fault, ball fault, and outer fault, and is classified into ten categories (one health state and nine fault states) according to different fault sizes. Moreover, CWRU consists of four motor loads corresponding to four operating speeds. The experimental platform of the CWRU dataset can be seen in Figure 4a.
- (2) *PU* [39]. PU bearing datasets were provided by Paderborn University (PU) for bearing fault diagnosis based on vibration signal. The PU dataset contains data for variable operating conditions. For example, different radial forces on the bearings, different load torques on the drive system, and different speeds of the drive system. The vibration signals are obtained at a sampling rate of 64 kHz under three types of bearings: realistically damaged bearings, artificially damaged bearings, and healthy bearings. Artificial damages arise in the inner race or outer race, and realistic damages arise in the form of pitting or plastic deformation. The experimental platform of the PU dataset can be seen in Figure 4b.
- (3) *IMS* [40]. The IMS dataset is from Prognostics Center Excellence through the prognostic data repository contributed by Intelligent Maintenance System (IMS), University of Cincinnati. The experiment performs bearing run-to-failure tests under constant loads on a specially designed test rig. The rotation speed is 2000 r/min and the sampling frequency is 20 kHz. The tests with outer race failure are investigated and data from different life cycle stages are considered, including severe failure, moderate failure, outer raceway incipient failure, and healthy condition.



(a) CWRU Dataset experimental platform.



(b) PU Dataset experimental platform.

Figure 4. Experimental platforms for CWRU dataset [38] and PU dataset [39]. Reprinted with permission from Ref. [38]. Copyright 2015 Elsevier.

3.1.2. Fault Diagnosis of Power Plant Thermal System

Background. Power units are developing towards large scale, integration, and complexity. The thermal system is one of the most critical components of a large-scale power unit [132]. In recent years, renewable energy sources such as solar energy and wind energy have been integrated into the grid [133]. Due to their stochastic and intermittent nature,

the task of frequency and peak regulation still needs to be performed by coal-fired power plants. This may lead to harsher operating conditions for the units and a higher probability of failure. Therefore, accurate fault diagnosis of thermal systems is an urgent problem that needs to be addressed.

Sensors. The monitoring of thermal power unit operating conditions includes many physical parameters, such as the temperature and pressure of high-pressure heaters, and the flow rate of condensate, etc. Therefore, the commonly used sensors in the fault diagnosis of the power plant thermal system are temperature sensors, pressure sensors, flow rate sensors, etc.

Domain Gap. In the fault diagnosis of the power plant thermal system, the domain gap comes from two main areas.

- (1) *Fault Severity.* For example, for feedwater leak faults, different leak volumes represent different levels of severity [60,61]. Although they are all fault data, the data collected from them differ in distribution.
- (2) *Different Load Conditions.* The distribution of the fault data collected by the power plant thermal system is different under different load conditions [60,61], which leads to the domain shift.

3.1.3. Diagnosis of Ball Screw Failure

Background. Ball screws are important mechanical devices capable of converting rotary motion into high-precision linear motion, and they are widely used in feed drive systems for machine tools [134]. Ball screws have important accuracy tolerances, and even slight degradation of the components can increase the operational risk. For example, the preload level of ball screws is carefully calibrated to maximize the expected life of the ball screw without affecting the repeatability of the motion. However, with use, a reduction in preload reduces the stiffness of the ball screw assembly and may eventually lead to a loss of positional accuracy. Therefore, it is important to accurately and reliably identify the health of the ball screw.

Sensors. During the operation of ball screws, vibration signals are often used to determine if a ball screw is faulty. Therefore, a common sensor used for ball screw fault diagnosis is the accelerometer sensor (see Figure 5).

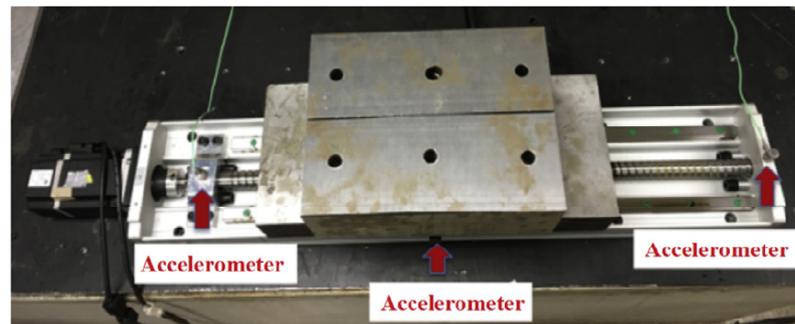


Figure 5. The ball screw test rig [63]. Reprinted with permission from Ref. [63]. Copyright 2020 Elsevier.

Domain Gap. In the fault diagnosis task of ball screws, the domain gap is mainly reflected in the differences in the data obtained by the sensors at different positions. Unlike rolling bearings, ball screw mechanisms have a more complicated motion trajectory in an operation cycle. Critical elements that need to be monitored, such as the ball screw nut assembly, are in motion when the ball screw is engaged. The ball screw nut also carries the work table on which workpieces are loaded, resulting in limited space available for sensor installation. One direction to solve this problem is to install the sensor on non-moving parts of the ball screw assembly, such as the front and end bearings, or in a location with proper space, such as the side of the table. However, due to the complex structure of the ball screw, the data collected by sensors placed in different locations vary greatly. Therefore, in the

diagnosis of ball screw failure, the domain gap is mainly reflected in the difference in data collected by sensors in different positions.

3.2. Transportation

This section presents deep unsupervised domain adaptation of relevant components in transportation. This includes the capacity estimation of lithium-ion batteries, remaining useful lifetime estimation of turbofan engines and bearings, and fault diagnosis of the gearbox.

3.2.1. Capacity Estimation of Lithium-Ion Batteries

Background. Due to their low self-discharge rate, low manufacturing cost, and high energy density [135], lithium-ion batteries have been widely deployed as energy storage devices in many fields, such as electronics and electric vehicles [136–139]. However, as the charge/discharge cycles increase, the performance of batteries will degrade due to the degradation of their electrochemical components. To ensure reliable operation and safety, it is important to estimate the capacity of individual cells online.

Sensors. The sensors commonly used in lithium-ion battery capacity estimation include current sensors, voltage sensors, and temperature sensors.

Domain Gap. In the actual operation of the battery system, different discharging and charging protocols will result in inconsistent capacity decay rates for different cells. This will result in significant differences in the data collected. In addition, due to individual differences in cell type or manufacture, batteries mostly show different degradation traces even under similar discharge/charge conditions.

Datasets. There are two commonly used public data sets for capacity estimation of lithium batteries, namely NASA Battery [64] and CALCE Battery [65].

- (1) *NASA Battery Dataset* [64]. The NASA Battery Dataset is a dataset of lithium-ion batteries published by the NASA Ames Prognostics Center of Excellence. These batteries are commercially available 18650 lithium-ion batteries and operate through three different operating profiles (impedance, discharge, and charge) at different ambient temperatures. Repeated charge and discharge cycles result in accelerated aging of the batteries. Data collection was stopped when the batteries reached the end-of-life criteria of 30% fade in rated capacity.
- (2) *CALCE Battery Dataset* [65]. The CALCE Battery dataset is a lithium-ion battery dataset presented by the Center for Advanced Life Cycle Engineering (CALCE) of the University of Maryland. These batteries undergo operations similar to NASA Battery, i.e., operation at different ambient temperatures through three different operating conditions (impedance, charging, and discharging). Each dataset contains five signals (i.e., discharge energy, internal resistance, voltage, current, and discharge time).

3.2.2. Remaining Useful Lifetime Estimation

Background. Remaining Useful Lifetime (RUL) relates to the amount of time left before a piece of equipment is considered not to perform its intended function [77,140,141]. Accurate RUL prognostics enable assessing the health status of equipment and planning future maintenance actions.

Sensors. Common applications in RUL estimation include RUL estimation of turbofan engines and RUL estimation of bearings. Commonly used sensors for the former are temperature sensors, pressure sensors, flow sensors, etc. The latter commonly used sensors include acceleration sensors and temperature sensors.

Domain Gap. In the task of RUL estimation, what leads to the domain gap are the operating conditions and failure modes.

- (1) *Different Failure Modes.* Taking RUL estimation of bearings as an example, the degradation of a bearing running to failure is usually associated with different failure modes, including inner failure, outer failure, etc., it is not a single failure behavior. Bearings in failure behavior under the action of the pressure and thermal strain usually lead to

other failure modes. The degradation of bearings exhibits different behaviors due to different failure behaviors and, therefore, may result in differences in the distribution of features.

- (2) *Different Operating Conditions*. In many practical applications, complete life degradation data are only available for bearings under certain operating conditions. However, the degradation process is often different when the operating conditions are different. Therefore, different operating conditions can also lead to a domain gap.

Datasets. Commonly used datasets in the estimation of RUL include the C-MAPPS dataset [68] and the IEEE PHM Challenge 2012 bearing dataset [69].

- (1) *C-MAPPS* [68]. Commercial Modular Aero-Propulsion System Simulation (C-MAPPS) datasets contain degraded data for turbofan engines (see Figure 6) and consist of four different datasets, each containing information from 21 sensors and 3 operational settings. Each dataset has several degraded engines, divided into training and test instances. The engines in the datasets start with various degrees of initial wear but are considered healthy at the beginning of each record. As the number of cycles increases, the engines begin to deteriorate until they can no longer function. At this point, the engines are considered unhealthy. Training datasets collect run-to-failure information throughout their life cycle until failure. Unlike the training datasets, the test datasets contain temporal data that terminates some time before the system fails.
- (2) *IEEE PHM Challenge 2012 bearing dataset* [69]. The IEEE PHM Challenge 2012 bearing dataset is collected from PRONOSTIA [69], which includes three key parts: the rotating part, load part, and data collection part. The platform provides the experimental run-to-failure data of rolling bearings through accelerated degradation experiments. Vibration and temperature signals are collected during all experiments. The frequency of vibration signal acquisition is 25.6 kHz. The frequency of temperature signal acquisition is 10 Hz. 600 samples are recorded each minute. The termination criterion of the experiment is set to a vibration amplitude of 20 g. The dataset is tested on 17 bearings under three different working conditions.

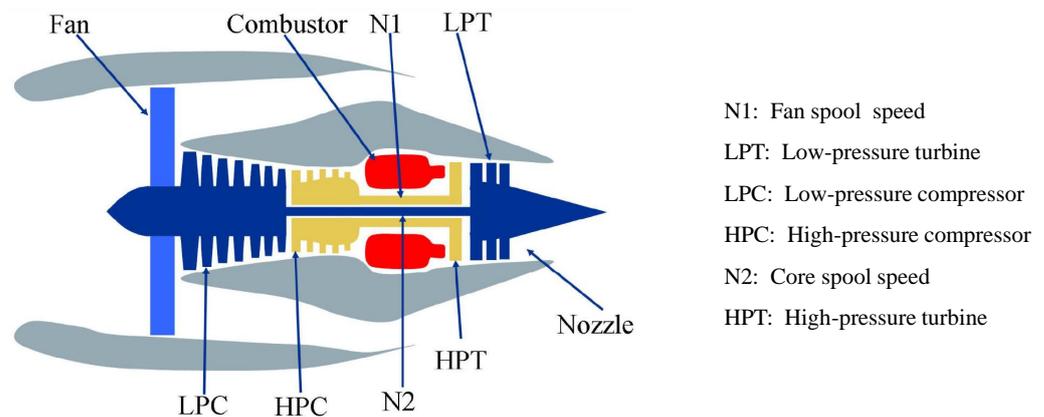


Figure 6. Diagram of the engines in C-MAPPS [68] dataset. Reprinted with permission from Ref. [68]. Copyright 2008 IEEE.

3.2.3. Gearbox Fault Diagnosis

Background. In the mechanical transmission system, the gearbox is a frequently used transmission type that has been widely used in wind turbine [142,143], helicopter [144,145], automobile [146,147], et al. However, due to the harsh working conditions and complex meshing mode, gearboxes are easily subject to gear faults. The failure of the gearbox often leads to the failure of the whole mechanical system and brings great losses. Therefore, an effective and accurate fault diagnosis of the gearbox is of great importance in the field of transportation.

Sensors. In fault diagnosis of the gearbox, accelerometers are commonly used to obtain the vibration signal of the gearbox. Usually, one accelerometer is fixed in the X-axis direction and one in the Y-axis direction to obtain vibration information in different directions.

Domain Gap. The domain gap in gearbox fault diagnosis comes from different operating conditions, including speed and load. This frequent change in speed and load will greatly affect the domain distribution.

3.3. Biosignal

This section presents two research directions of UDA in biosignal, namely EEG-based brain-computer interface (BCI), human activity recognition (HAR), EMG-based muscle-computer interface (MCI), and gait analysis.

3.3.1. EEG Based BCI

Background. Brain-computer interface is an important issue in biosignal and it enables users to communicate directly with computers using brain signals [97,98,148–150]. There are three general types of BCI, non-invasive BCI, invasive BCI, and partially invasive BCI. Electroencephalography (EEG) is an electrophysiological non-invasive technique for recording the electrical activity generated by the human brain, and it is increasingly being used in BCI tasks due to its safety, low cost and convenience.

Sensors. EEG signals usually are collected using a special device called an electroencephalogram. This device uses the principle of differential amplification or recording voltage differences between different points using a pair of special metal plate electrodes that compares one active exploring electrode site with another neighboring or distant reference electrode.

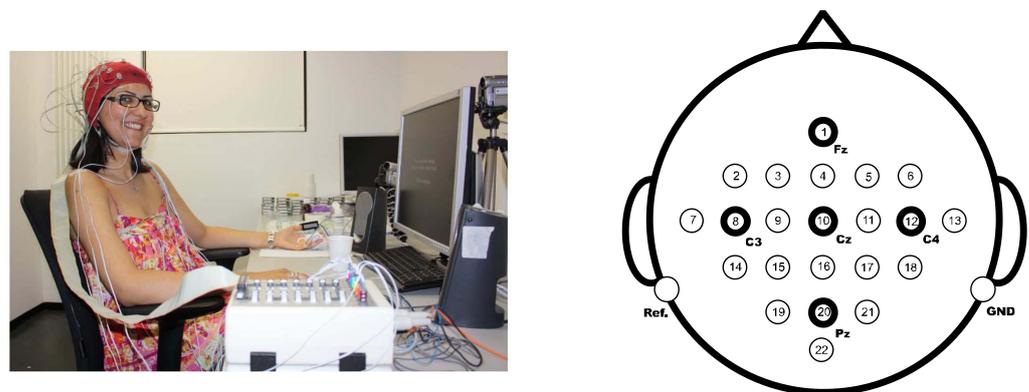
Domain Gap. The domain gap of EEG-based BCI is mainly reflected in different subjects and different sessions.

- (1) *Different Subjects.* Brain signals show high variability among subjects due to inherent background neural activities, concentration levels, fatigue, etc. In this regard, deep learning-based BCI faces a great challenge in that classifiers trained on one subject cannot be directly used to decode brain signals from other subjects.
- (2) *Different Sessions.* EEG signals are weak, susceptible to interference and noise contamination, nonstationary for the same subject, and varying across different sessions. Therefore, there are also differences in the distribution of signals collected at different sessions on the same subject.

Datasets. The common applications of EEG-based BCI are emotion recognition and motor imagery. The datasets related to emotion recognition include SEED [87] and DEAP [88]. The datasets related to motor imagery include BCI Competition IV-IIa [89] and BCI Competition IV-IIb [151].

- (1) *SEED* [87]. Shanghai JiaoTong University Emotion EEG Dataset (SEED) dataset consists of 15 participants, each of whom was required to watch 15 Chinese movie clips to induce three different kinds of emotions: negative, neutral, and positive. All of them are native Chinese with 7 males and 8 females. The movie clips are carefully selected as the stimuli for the experiment. All movies need to be comprehensible and neither too long nor too short to elicit sufficient emotional meaning. The EEG signals were recorded by an ESI NeuroScan System with a 62 electrode cap at a sampling rate of 1000 Hz.
- (2) *DEAP* [88]. The Database for Emotion Analysis using Physiological Signals (DEAP) consisted of 32 subjects. Each subject was exposed to 40 1-min long music videos as emotional stimuli, while physiological signals were recorded. There were 40 EEG trials recorded for each subject, each corresponding to an emotion elicited by a music video. After watching each video, subjects were asked to assess the emotion they really felt in five ways: (1) familiarity (associated with the knowledge of the stimulus); (2) liking (associated with preference); (3) dominance (associated with control power); (4) arousal

- (associated with excitation level); and (5) valence (associated with pleasantness level). The scale ranges from 1 (weakest) to 9 (strongest). The EEG signals were recorded by Biosemi Active Two devices at a sampling rate of 512 Hz and downsampled to 128 Hz.
- (3) *BCI Competition IV-IIa* [89]. The BCI Competition IV-IIa dataset (see Figure 7) consists of EEG signals of four different motor imagery tasks (left hand, right hand, foot, and tongue), which were acquired from nine subjects. According to the 10–20 system, all signals were recorded from 22 Ag/AgCl electrodes. The signals were sampled at 250 Hz and band-pass filtered between 0.5 Hz and 100 Hz.
 - (4) *BCI Competition IV-IIb* [151]. The BCI Competition IV-IIb dataset involves left-hand and right-hand motor imagery activity and contains three bipolar channel EEG signals from nine subjects. Each subject was sampled at a rate of 250 Hz and five time periods were collected.



(a) A participant in an EEG-based emotion recognition experiment.

(b) Electrode montage corresponding to the international 10-20 system.

Figure 7. A participant in an EEG-based emotion recognition experiment [88] and electrode montage of BCI Competition IV-IIa [89]. Reprinted with permission from Ref. [88]. Copyright 2011 IEEE.

3.3.2. Human Activity Recognition

Background. Human activity recognition (HAR) is the recognition of user behavior from observed data collected by a set of sensors [152–154]. For example, inferring whether a user is standing or lying down based on data collected by accelerometers and gyroscopes embedded in a smartphone. HAR can be used for fall detection in the elderly, lifelogging systems for monitoring energy consumption, and digital assistants for weight lifting. As smart sensing technologies become more prevalent, more and more HAR systems are being deployed in our living environments. Therefore, accurate human activity recognition is of great importance to the biosignal field.

Sensors. In the task of human activity recognition, commonly used sensors are accelerometers and gyroscopes. There are also some recent studies [105,106] using wifi sensors for activity recognition.

Domain Gap. In the field of human activity recognition, domain gaps exist mainly in the differences between sensor signals acquired by different body parts of the same user and differences in sensor signals acquired by different users. Besides, there are also some differences in the signals acquired by different sensors, which can also lead to the domain shift of the model.

- (1) *Different Body Parts.* Users often change the position of a sensor or wearable device based on their preferences and current activity. Despite doing the same activity, the sensor signals obtained from different body parts are very different. Therefore, transferring the activity model learned at one body location to another body location can reduce annotation costs and improve recognition accuracy and robustness to wearable device position changes.

- (2) *Different Users*. Even when doing the same activity, there is some variation in the sensor signals collected by different users. A HAR system has a large number of users, and it is impractical or infeasible to collect enough ground truth from each user to build a model of the activity. Therefore, it is also very important to reduce the domain shift of the model between different users.
- (3) *Different Sensors*. There are some differences in the data collected by different sensors (e.g., smartphones and smartwatches). Therefore, how to improve the performance of the model on newly deployed sensors without collecting any activity labels for new sensors is also a very important issue.
- (4) *Different Environments*. When using WIFI sensors to identify human activities, different WIFI environments often result in biased data collection. For example, different indoor structures have different signal strengths and distributions of WIFI reflections [105]. Another example is that in in-vehicle human activity behavior recognition, different speeds of the car and interference from different wireless signals outside the car can lead to significant bias in the collected data [106].

Datasets. Common datasets in the field of human activity recognition include Opportunity [100], HHAR [101], PAMAP2 [102], and RealWorld [99].

- (1) *Opportunity* [100]. The Opportunity Activity Recognition dataset includes data recorded from 4 participants performing 5 activities: lying, sitting, walking, standing, and null. Users use custom-designed sports jackets and shoes to collect accelerometer signals from a total of 19 body positions. Accelerometer data were sampled simultaneously from the device at a sampling rate of 30 Hz. Users were asked to perform predefined daily home activities for 15–25 min. This is a challenging dataset since the sensors were placed on clothing rather than being tied to specific body locations.
- (2) *HHAR* [101]. The Heterogeneity Human Activity Recognition (HHAR) dataset includes data recorded from 9 participants performing 6 activities: walk, stand, stairs up, stairs down, sit, and bike. Users were instrumented with 4 smartwatches and 8 smartphones. All 4 smartwatches were worn on each arm, and smartphones were placed around the waist. Accelerometer data were sampled simultaneously from the device at a sampling rate of 50 to 200 Hz. Each activity was performed for 5 min for each user. This dataset is noisy compared to the other datasets because the timestamps are not continuous and the sampling rate is not stable.
- (3) *PAMAP2* [102]. The physical activity monitoring for aging people dataset 2 (PAMAP2) includes data recorded from 9 subjects performing 18 different activities, including ascending stairs, cycling, descending stairs, etc. Users used 3 wireless inertial measurement devices placed at 3 different body locations: head, chest, and ankle. Accelerometer and gyroscope data were sampled simultaneously from these devices at a sampling rate of 100 Hz. Each user performed up to 3 min of each activity.
- (4) *RealWorld* [99]. The RealWorld HAR dataset includes data recorded from 15 participants performing 7 activities: walking, running/jogging, sitting, standing, climbing stairs down and up, lying, and jumping. The user places the smartphone and smartwatch in seven different body positions: head, chest, upper arm, waist, forearm, thigh, and calf. Accelerometer and gyroscope data are sampled simultaneously from the device at a sampling rate of 50 Hz. Figure 8 shows the location of the sensors in the different datasets.

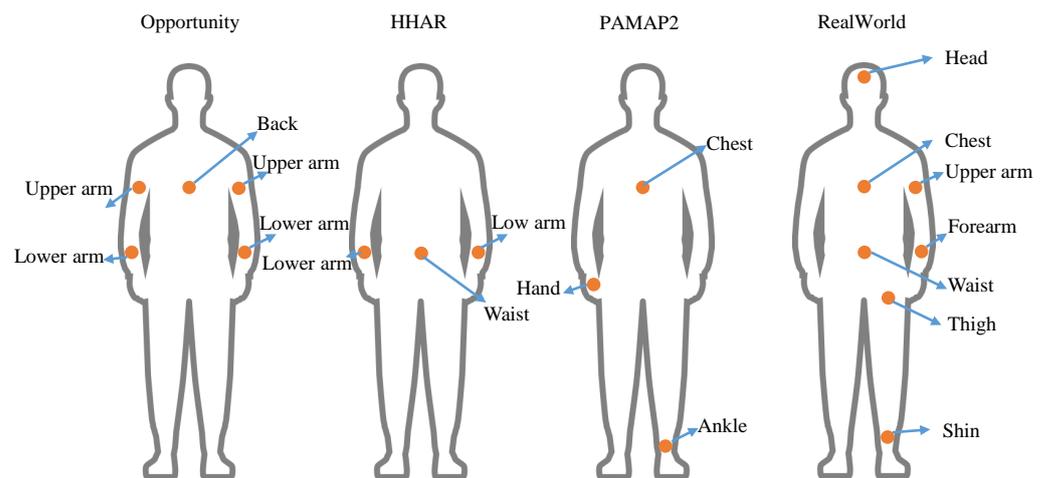


Figure 8. Placement of sensors in the Opportunity dataset [100], HHAR dataset [101], PAMAP2 dataset [102] and RealWorld dataset [99].

3.3.3. EMG Based MCI

Background. The muscle-computer interface (MCI) is an interaction method that translates myoelectrical signals directly from a mere reflection of muscle activity to an interactive command that conveys the intent of the user’s movement. It does not depend on actions performed by the user on a physical device, nor on externally visible or audible actions, thus enabling a readily available input mechanism in myoelectric control. As the technical core of non-invasive MCI, electromyography (EMG) technology is a technique that uses one or more electrodes to measure the electrical activity of muscles from the skin surface. Due to its safety and convenience, it is increasingly being used for MCI tasks.

Sensors. EMG signals are usually collected using a special device called an electromyogram. When muscles contract, they release a burst of electrical activity or electrical impulses that circulate through adjoining bones and tissues. The myoelectric potential is detected with surface and reference electrodes, and this potential is amplified using differential amplifiers to detect muscle or limb activation.

Domain Gap. The domain gap of EMG-based MCI is mainly reflected in different sessions and different subjects.

- (1) *Different Sessions.* In practice, the electrode position may change when the EMG is taken off or put on, sweating, etc. This change is mainly caused by changes in electrode conductivity, electrophysiological changes, and electrode movement. This inherent non-stationarity leads to great variability in the signals collected by the EMG.
- (2) *Different Subjects.* Since the EMG signal is a biological signal, the obtained signal varies considerably between subjects due to their physical condition. In this case, the variability of the data comes from the differences in the human body.

Datasets. Common datasets for EMG-based MCI include the CapgMyo dataset [111], the NinaPro dataset [112] and the CSL-HDEMG dataset [113].

- (1) *CapgMyo* [111]. CapgMyo is an EMG-based gesture recognition dataset. It consists of 128 channels of HD EMG data acquired from 23 intact subjects with a sampling rate of 1 KHz. It consists of 3 sub-databases, DB-a, DB-b, and DB-c. DB-a contains 8 isometric, isotonic hand gestures obtained from 18 subjects. DB-b contains 8 isometric, isotonic hand gestures from 10 subjects in two recording sessions on different days. DB-c contains 12 basic movements of the fingers obtained from 10 subjects. Figure 9 illustrates the acquisition setup for the CapgMyo dataset.
- (2) *NinaPro* [112]. The Ninapro (Non-Invasive Adaptive Prosthetics) dataset is a representative publicly available dataset in the field of surface myoelectricity research. Ninapro contains a total of 10 sub-datasets. Here, the DB-1 sub-dataset is introduced as an example. DB-1 sub-dataset is used to develop the hand prostheses and contains sparse

multi-channel EMG recordings. It consisted of a total of 52 gestures performed by 27 intact subjects. Data are recorded at a sampling rate of 100 Hz, using 10 sparse electrodes placed on the subject's upper forearm.

- (3) *CSL-HDEMG* [113]. The CSL-HDEMG benchmark database is created specifically for EMG-based gesture recognition. EMG signals in CSL-HDEMG are recorded by using an electrode array with 192 electrodes covering the upper forearm muscles of five subjects performing a total of 27 gestures. The number of subjects in the CSL-HDEMG was relatively small compared to the CapgMyo dataset and the NinaPro dataset.

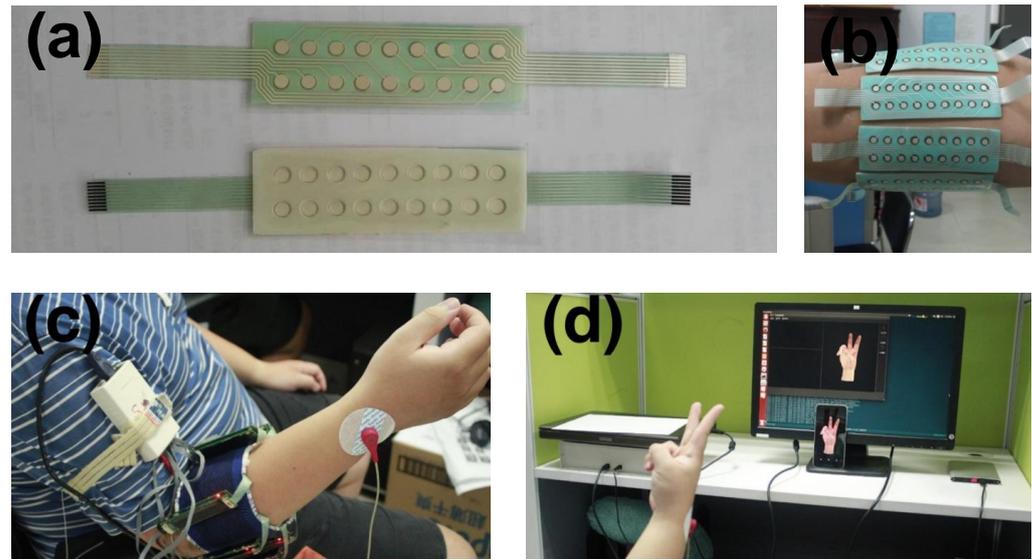


Figure 9. The acquisition setting-up for CapgMyo dataset [111]: (a) The EMG electrode array; (b) 8 EMG electrode arrays on the right forearm; (c) The EMG acquisition device ready for capture; (d) The software subsystem to present the guided hand gesture and record EMG data simultaneously.

3.3.4. Gait Analysis

Background. Walking and running are the most frequent movements in human life. Identifying gait phases is useful in a variety of applications. For example, physical therapists analyze gait phases to identify abnormal walking patterns. In sports science, gait phases are used to improve the skills and physical condition of athletes. Gait phases can also be used to synchronize the assistive forces of a wearable robot with the user's movements.

Sensors. The sensors used for gait analysis include IMU sensors (including accelerometers and gyroscopes), infrared (IR) sensors, radar sensors, cameras, and EMG electrodes.

Domain Gap. For gait analysis, the domain gap mainly comes from the different positions of sensor placement, different subjects, and different moving states.

- (1) *Different Positions of Sensors.* Such a problem is common in IMU sensor-based gait analysis. The position of the IMU may drift due to motion or personal preference, and it is difficult to precisely control the fixture of the IMU on users in real-world settings. Since IMU signals are location sensitive, the signals recorded at different locations can vary dramatically.
- (2) *Different Subjects.* Human gait is highly individualized; each individual has his or her own unique gait pattern. Therefore, the sources of domain gaps in gait analysis also include different subjects.
- (3) *Different Moving States.* In radar sensor-based gait analysis, differences in domains often stem from different states of motion. This is because when the test subject changes his or her motion state (different clothing, different motion speed, whether carrying an object, etc.), the radar echo intensity, Doppler bandwidth, or motion period will change. This also means that the distribution of gait data will also be shifted.

Datasets. Common gait analysis datasets include the Daphnet dataset [118], the OU-ISIR dataset [119], the CASIA-B dataset [120], and the CASIA-C dataset [121].

- (1) *Daphnet* [118]. This dataset was designed to benchmark methods for automatically identifying gait freezes from wearable accelerometers placed on the legs and hips. The data was collected from 10 subjects with Parkinson’s disease using three wearable accelerometers. The three sensors were placed on the ankle (shank), thigh above the knee, and hip. Users performed three tasks: walking in a straight line, walking with multiple turns, and finally a more realistic activity of daily living task.
- (2) *OU-ISIR* [119]. The OU-ISIR gait database was collected by the Institute of Scientific and Industrial Research (ISIR), Osaka University (OU). The dataset consists of people walking on the ground surrounded by 2 cameras at 30 fps and 640×480 pixels. These datasets are essentially distributed as silhouette sequences registered and size-normalized at 88×128 pixels.
- (3) *CASIA-B* [120]. The CASIA-B dataset is published by the Institute of Automation, Chinese Academy of Sciences (CASIA). It contains 124 individuals captured from 11 viewpoints ($0^\circ, 18^\circ, \dots, 180^\circ$). Each person walks six times in normal conditions, twice in their coats, and twice with bags to obtain a total of ten gait sequences. Overall, the CASIA-B dataset contains 13,640 sequences.
- (4) *CASIA-C* [121]. The CASIA-C contains 1530 video clips, which are captured with a thermal infrared camera with a resolution of 320×240 pixels and a frame rate of 25 frames per second. Each video clip depicts a person walking in one of four different ways, including walking normally (normal), walking fast (fast), walking slowly (slow), and walking with a backpack (backpack).

4. Methods

In this section, unsupervised domain adaptation methods for time series sensor data are investigated. The most commonly used methods include adaptation in the input space, adaptation in the feature space, adaptation in the output space, and model-based adaptation (see Table 2).

In order to describe the different UDA methods more clearly, we have predefined some notations. Suppose the source domain \mathcal{D}_s is defined as follows:

$$\mathcal{D}_s = \{(\mathbf{x}_i^s, \mathbf{y}_i^s)\}_{i=1}^{n_s} \quad \mathbf{x}_i^s \in X_s, \quad \mathbf{y}_i^s \in Y_s, \quad (2)$$

where \mathbf{x}_i^s is the i -th sample as defined in Equation (1). X_s is the union of all samples. \mathbf{y}_i^s is the i -th label of the i -th sample. Y_s is the union of all different labels, and n_s means the total number of source samples. The target domain is defined as :

$$\mathcal{D}_t = \{(\mathbf{x}_i^t)\}_{i=1}^{n_t} \quad \mathbf{x}_i^t \in X_t, \quad (3)$$

where \mathbf{x}_i^t is the i -th sample. X_t is the union of all samples. n_t means the total number of target samples. The goal of unsupervised domain adaptation is to use the labeled source domain data \mathcal{D}_s and the unlabeled target domain data \mathcal{D}_t to improve the representation of the deep model in the target domain samples.

Table 2. The classification results based on different domain adaptation methods.

Different Adaptation Methods	Description	References
Input space	By generating source domain samples that are very similar to the target domain, the gap in the source domain is reduced by supervised training.	[48,104,109]
Feature Space (mapping-based)	The features are mapped to some space and then a metric is used to reduce the distance between the source and target domains.	[41,44,46,47,51–54,59–63,66,67,70,71,73,76,78,79,81–84,86,91,103,104,107,122,123,127]
Feature Space (adversarial-based)	The discriminator is used to identify whether the generated features are from the source or target domain. The feature extraction network tries to fool the discriminator. As a result, the network is able to generate similar features for both source and target domain samples.	[42,43,45,49,50,55,67,71,72,74–77,80,81,84,85,85,92–96,105,106,108,124,126]
Output space	The labels with high confidence are selected for the target domain, and these pseudo-labels are used to do supervised training on the target domain samples.	[56–58,114]
Model-based	By constraining the parameters of the model, the model can be adapted to the sample of the target domain.	[91,110,111,115–117,125]

4.1. Adaptation in Input Space

Domain adaptation in the input space involves generating source domain samples that are very similar to the target domain, in the form of using generative adversarial networks or using prior knowledge. For example, in human behavior recognition, Sanabria et al. [104] use bi-directional generative adversarial networks [155] to generate source domain samples that are very similar to the target domain samples. Since the source domain samples are labeled, the performed transformed samples can be trained at scale using supervised learning, thus reducing the domain differences between the source and target domains. Another domain adaptation approach in the input space is to synthesize data artificially using a priori knowledge. For example, in fault diagnosis of bearings, Wang et al. [48] selects healthy real samples as the base signal for the synthesis process and uses expert knowledge to inject failure modes. Since the base signal encodes information about the operating and environmental conditions, the generated signal can be better adapted to the target domain, thus reducing the domain discrepancy.

4.2. Adaptation in Feature Space

Some domain adaptation methods unify the source and target domains by creating a domain-invariant feature representation, usually in the form of a feature extractor neural network. A feature representation is a domain invariant if the features follow the same distribution regardless of whether the input data is from the source or target domain. If a classifier can be trained to perform well on the source data using domain-invariant features, then the classifier generalizes well to the target domain because the features of the target data match the features for which the classifier is trained. Feature space adaptation methods typically include mapping-based methods and adversarial learning-based methods (see Figure 10). The next subsections will describe these two methods in detail.

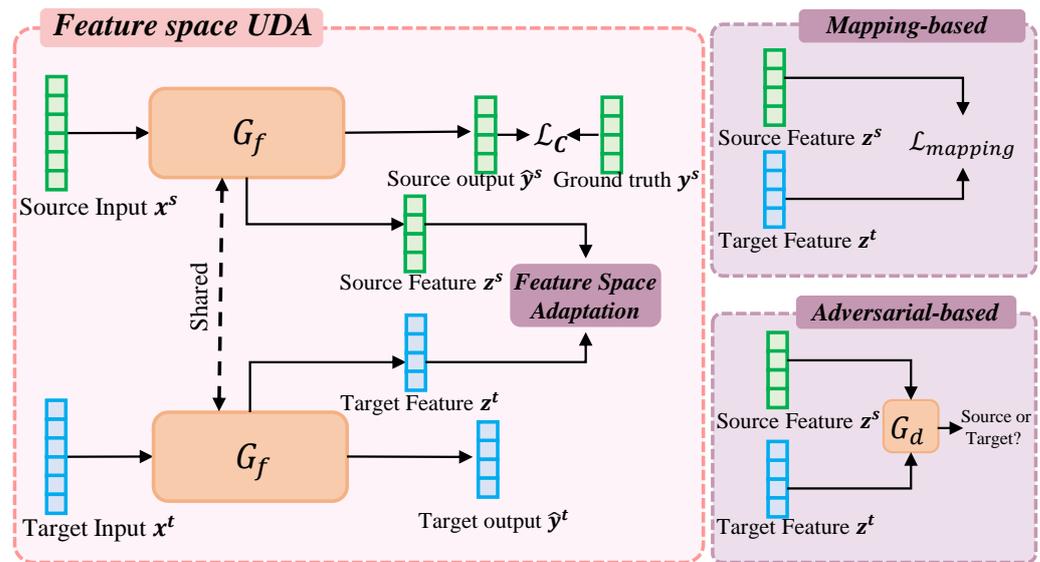


Figure 10. Diagram of feature space adaptation, including mapping-based (top right) and adversarial-based (bottom right) approaches.

4.2.1. Mapping-Based Methods

The mapping-based approach includes mapping instances of the source and target domains to the feature space by a feature extractor and reducing the distance of the mapped features. Depending on the metrics, feature alignment methods include maximum mean discrepancy (MMD), multi-kernel MMD (MK-MMD), correlation alignment (CORAL), etc.

MMD. Maximum mean discrepancy is a two-sample statistical test of the hypothesis that two distributions are equally based on observed samples from the two distributions. The test is computed from the difference between the mean values of a smooth function on the samples from two domains. If the means are different, then the samples are most likely not from the same distribution. The smooth functions chosen for MMD are unit balls in feature reproducing kernel Hilbert spaces (RKHS) since it can be proven that the population MMD is zero if and only if the two distributions are equal. Assuming that the marginal distributions of the source and target domains are $P(X_s)$ and $P(X_t)$. The features extracted from the source domain sample x^s and the target domain sample x^t after passing through the network G_f are z^s and z^t , respectively. The naive MMD can be expressed as

$$\mathcal{L}_{\text{mapping}} = \|\mathbb{E}_P(\phi(z^s)) - \mathbb{E}_Q(\phi(z^t))\|_{\mathcal{H}_k}^2, \quad (4)$$

where \mathcal{H}_k is RKHS using the kernel k (the Gaussian kernel is often used as the kernel) and $\phi(\cdot)$ is the mapping to RKHS. Equation (4) makes the network extract similar features for the source and target domains, thus reducing the domain gap between them.

Suppose the predicted values of the source domain sample x^s and the target domain sample x^t after passing through the network G_f are \hat{y}_i^s and \hat{y}_i^t . The supervised learning loss of the source domain sample can be expressed:

$$\mathcal{L}_c = -\mathbb{E}_{(x_i^s, y_i^s) \in \mathcal{D}_s} \sum_{c=0}^{C-1} \mathbf{1}_{[y_i^s=c]} \times \log \hat{y}_i^s, \quad (5)$$

where C is the number of classes. Combining the Equation (4), the optimization objective of the network G_f is

$$\arg \min_{\theta_f} \mathcal{L}_c + \lambda_{\text{mapping}} \mathcal{L}_{\text{mapping}}, \quad (6)$$

where θ_f is the parameter of the network G_f . λ_{mapping} is the balance parameter.

MK-MMD. The choice of parameters for each kernel is crucial to the final performance. Since different kernels have different performances, in practice, it is difficult to know which one to choose. To solve this problem, many scholars [57,70,78,81] have proposed to use MK-MMD, i.e., to build a total kernel with multiple kernels. It can maximize the testing power of two samples. For example, in the remaining useful lifetime estimation, Zhuang et al. [70] use MK-MMD to make the source domain features and the target domain features similar. In the fault diagnosis of bearings, Yang et al. [57] uses MK-MMD to reduce the difference between source and target domains among multiple features output from the network.

CORAL. The CORAL [60,84] is similar to MMD with a polynomial kernel and is computed from the distance between the second-order statistics (covariance) of the source and target features. For domain adaptation, the alignment component consists of computing the CORAL loss between the outputs of the two feature extractors. The CORAL loss function has applications in a wide range of industries. For example, Qin et al. [84] uses CORAL loss to improve the fault diagnostic capability of planetary gearboxes under different operating conditions. In fault diagnosis of power plant systems, Wang et al. [60] used CORAL loss to reduce the domain shift of the trained model.

4.2.2. Adversarial-Based Methods

Adversarial-based methods refer to an adversarial approach that uses domain discriminators to reduce the difference in feature distribution between the source and target domains generated by the feature extractor. Like MMD and MK-MMD, adversarial-based methods are defined to solve the problem of unequal distributions. Unlike MMD and MK-MMD, adversarial-based methods do not use some metric to measure the distance between features directly, but introduce an additional discriminator G_d (see Figure 10). The source and target domain features extracted by the feature extractor are input to the discriminator separately, and the parameters of the discriminator are optimized so that the discriminator can distinguish between them. The parameters of the feature extractor are optimized so that the discriminator is not able to distinguish between them. The formula is expressed as follows

$$\mathcal{L}_{adv} = \mathbb{E}_{\mathbf{x}_i^s \in \mathcal{D}_s} \log[G_d(\mathbf{z}_i^s)] + \mathbb{E}_{\mathbf{x}_i^t \in \mathcal{D}_t} \log[1 - G_d(\mathbf{z}_i^t)]. \quad (7)$$

Combining the supervised learning loss of the source domain samples in Equation (5), the optimization objective of the network G_f can be expressed as

$$\arg \min_{\theta_f} \mathcal{L}_c + \lambda_{adv} \mathcal{L}_{adv}, \quad (8)$$

where λ_{adv} is the balance parameter. The objective optimization of discriminator G_d can be expressed as

$$\arg \max_{\theta_d} \mathcal{L}_{adv}. \quad (9)$$

By alternately optimizing Equations (8) and (9), the classification network G_f is able to extract domain-invariant features, thus reducing the domain gap between the source and target domains.

Like mapping-based methods, there are a large number of methods in UDA that are based on adversarial learning. For example, Wang et al. [106] uses adversarial learning to improve the model's ability to recognize human behavior. Zhang et al. [83] on the other hand, uses multilayer adversarial methods to improve the fault diagnosis of gearboxes. In the capacity estimation of lithium-ion batteries, Ye et al. [67] combines the strategies of MMD and adversarial learning to improve the performance of the model in the target domain.

4.3. Adaptation in Output Space

Pseudo-label-based domain adaptation methods [56–58] are the most common domain adaptation methods in the output space. Using classification as an example, the model’s confidence in different categories can be determined by looking at the softmax distribution in the last layer: uniformity indicates uncertainty, while a much higher probability for one category than the others indicates higher confidence. Applying this to domain adaptation, a diverse ensemble trained on the source data can be used to label the target data. These now-labeled target examples can then be used to train a classifier on the target data if the ensemble has a high level of confidence. This is thus a supervised process in the target domain, i.e., the network is trained using the assumed true labels (pseudo labels).

Pseudo-label-based domain adaptation methods also have some applications in solving the domain gap problem of time series sensor data. For example, Song et al. [56] uses a pseudo-label-based approach to do fault diagnosis of gearboxes. Yang et al. [57], on the other hand, combines MK-MMD and pseudo label to enhance the representation of the model under different operating conditions. Zhang et al. [58] proposes an iterative matching network enhanced with a selective sample reuse strategy. They use specially designed filters to select pseudo-label signals to increase the proportion of correctly labeled signals in the iterations and improve the performance of the model in the target domain.

4.4. Model-Based Adaptation Methods

The model-based domain adaptation approach is to reduce the domain shift of a model by imposing constraints on its parameters. For example, Jimenez et al. [91] use adaptive batch normalization [156] to align the domain distributions in terms of their first and second statistical moments. Khan et al. [110] proposes a transitional adaptation learning model specifically tuned to the properties of CNNs. They assume that the weight distribution in the convolutional layers remains essentially constant across contexts, and thus automatically adjusts the weights while minimizing the divergence of the weight distribution. Since the designed transfer learning method implicitly labels the activity of the source context, it does not require any explicitly labeled target training data to improve the performance of the model in the target domain.

5. Settings

This section describes the differences in adaptation settings caused by different numbers of domains in unsupervised domain adaptation, i.e., single-source single-target domain adaptation, multi-source single-target domain adaptation, and single-source multi-target domain adaptation. The differences between the different settings are shown in Figure 11 and Table 3. Each of the adaptation settings is described in detail next.

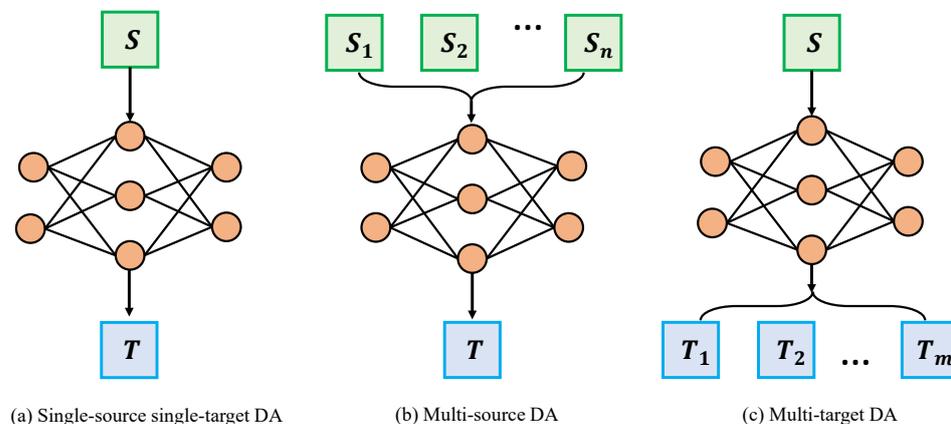


Figure 11. Different adaptation settings. S and T denote the source and target domains, respectively. S_1, S_2, \dots, S_n and T_1, T_2, \dots, T_m denote n different source domains and m different target domains, respectively.

Table 3. Classification results based on the number of source and target domains.

Different Settings of Domain Adaptation	Advantages	References
Single-source single-target domain adaptation	This setup is simple and more focused on the target area.	[41,42,46–49,51,53–63,66,67,70–86,91–93,95,96,103,104,106,107,109,110,114–116,123,125,126]
Multi-source domain adaptation	Each source domain has its own focus and can integrate different aspects of information.	[43–45,52,94,105,108,111,117,124,127]
Multi-target domain adaptation	The trained model can be adapted to multiple working conditions simultaneously.	[50,122]

5.1. Single-Source Single-Target Domain Adaptation

The vast majority of current domain adaptation algorithms focus on a single-source single-target setup. This setup is relatively simple, assuming only one source and one target domain. The single domain gap motivates a large number of approaches, including domain adaptation in the input space, domain adaptation in the feature space, etc. However, this single-source single-target setup can have limitations. In reality, the sensor deployment environment is always changing, and models adapted in a single target domain are difficult to transfer to other target domains (data collected under changing sensor deployment conditions). Moreover, during domain adaptation, the available supervised samples may come from multiple source domains with different distributions. Therefore, it is also important to utilize multiple source domains with labels to improve the performance of the target domain.

5.2. Multi-Source Domain Adaptation

Unlike single-source domain adaptation, multi-source domain adaptation can obtain more comprehensive knowledge by fusing supervised samples from multiple source domains and mitigate the risk of overfitting problems by having more training samples, which is beneficial to improving the performance of the target domain.

In order to effectively fuse different source domains, Zhu et al. [45] proposes a multiple adversarial learning strategy with multiple domain discriminators to learn domain invariant and discriminative features of the task. Through the multi-adversarial learning strategy, the samples from different rolling bearings are projected into a shared feature subspace so that domain invariant features are obtained. Meanwhile, the extracted features are also task-discriminative since sufficient labeled samples in the source domains are available. Xia et al. [41], on the other hand, proposes a novel multi-source domain adaptation model, which uses a feature learner to generate features for each source and target domain data to enable a joint weight classifier to predict the target labels. They also introduce an MMD-based distance metric to reduce the distance between all source and target domains. In the training process of the model, an intra-class matching training strategy is used to match the distribution of each domain to improve the recognition accuracy of the target domain. In the field of human activity recognition, Chakma et al. [108] proposes a deep multi-source adversarial domain adaptation framework that opportunistically helps select the most relevant feature representations from multiple source domains and establish such mappings to the target domain by learning the perplexity scores.

Although multi-source domain adaptation achieves better results compared to single source domain adaptation, this setting may lead to negative adaptation, i.e., some source domains may not have a positive effect on improving the model's performance in the target domain. To address this issue, Wei et al. [55] proposes a multi-source domain adaptation model for weighting different source domains. In particular, they insert a weighting

module in the output space to avoid negative adaptation of source domains by adaptively controlling the weights of different source domains.

5.3. Multi-Target Domain Adaptation

In practical application scenarios, different working conditions and different environments result in different domain gaps. Therefore, it is important to train a model to have a good performance on multiple target domains at the same time. There are two naive ways of directly extending domain-specialized UDA to work on multiple target domains, that are (1) training a single model on combined data from multiple target domains, and (2) training multiple models individually for each target domain. Unfortunately, these methods are not appropriate to handle multi-target domain adaptation problems because they would suffer from performance degradation due to the mismatching of multi-target domains [157].

How to explore the potential connection between multiple target domains is an important research topic. There is very limited research on multi-target-based domain adaptation in time series sensor data. Ref. [50] proposes a framework for single-source to multi-target domains for bearing fault diagnosis. They first train the source feature extractor to obtain class discriminative features using the labeled source domain. Then, the target feature extractors are initialized by the weights of the source feature extractor and, thus, inherit the class-discriminative property. On the other hand, a discriminator network is trained to distinguish between the source and multi-target features. To obtain domain-invariant features among different targets, They adversarially update multi-target feature extractors to generate features that can be indistinguishable for the discriminator. During testing, the scalable model can take any of the target domains and generate source-like features, where the trained source classifier is able to generalize well to any of the targets.

6. Discussion

As shown in Section 3, UDA technology is heavily used in a variety of industries due to the complex environment in the real world. Nevertheless, there are still some potential research directions worth exploring. For example, with the popularity of smart wearable devices, more and more manufacturers are using smart devices to obtain the user's heart rate and electrocardiogram to measure the user's physical condition. However, models trained on some people do not generalize well to others due to the differences in fitness of different people. Therefore, how to reduce the domain shift of heart rate recognition models or electrocardiogram recognition models is a problem worth investigating.

In addition to the applications of UDA, there will be some room for exploration in the UDA approach. Most of the deep models used by existing methods are CNN and RNN, and if the backbone network is replaced with Transformer [30] that deals with time series sensor data, it may also improve the adaptation performance. Another way is to use knowledge distillation [158] to train the teacher in the source domain, and combine it with pseudo label generation to teach the student in the target domain, and improve the performance through continuous iteration. Similarly, for the multi-source domain adaptation approach, multiple teachers can be used, each with a different direction of focus. In this way, it may be possible to better transfer models trained with data from different source domains to the target domain, thus improving the performance in the target domain. Another strategy that can be explored for unsupervised domain adaptation is to use contrast learning [159]. That is, a large amount of unlabeled data is used to enhance the ability of the model to acquire features through contrast learning, and then traditional unsupervised domain adaptation methods are used to make the features in the source and target domains close in distribution, and thus reduce the domain differences between the source and target domains.

In terms of UDA settings, existing methods rarely explore multi-source to multi-target domain adaptation. This adaptation setting would include the advantages of both multi-source adaptation and multi-target adaptation, i.e., the source domain contains data collected under multiple conditions and can simultaneously contribute to improving the

model performance in the target domain. Also, since multiple target domains are included, the adapted model can be generalized to multiple operating conditions.

7. Conclusions

Complex environments and variable working conditions result in deep learning models trained with time series sensor data collected in one condition performing poorly when tested in another. In recent years, more and more scholars have adopted the unsupervised domain adaptation approach to solve the domain shift problem of deep models. This survey systematically studied, analyzed and summarized the unsupervised domain adaptation technique based on time series sensor data from three aspects: application, method, and setup. In particular, this survey analyzed and summarized which sensors are used in different application directions, where the domain gaps are, and what data are publicly available. Moreover, this survey classified and compared different unsupervised domain adaptation methods and discussed and summarized the advantages of various settings. This survey also provided an outlook on potential application areas, explored feasible research directions, and discussed additional settings for unsupervised domain adaptation. Through this survey, we hope to provide a comprehensive and systematic understanding to readers who are investigating the direction of unsupervised domain adaptation with time series sensor data.

Author Contributions: This paper is a collaborative work by all authors. Conceptualization, X.Y.; investigation, X.Y. and Y.S.; writing-original draft preparation, Y.S.; writing-review and editing, X.Y., Y.S. and J.Y.; supervision, X.Y. All authors have read and agreed to the submitted version of the manuscript.

Funding: This work was supported in part by the National Key R&D Program of China Grand No. 2020YFB1708002, and NNSFC Grant No. 61971008.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: No new data were created or analyzed in this study. Data sharing is not applicable to this article.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

CNN	Convolutional Neural Network
RNN	Recurrent Neural Network
LSTM	Long Short-Term Memory
GRU	Gated Recurrent Unit
UDA	Unsupervised Domain Adaptation
CWRU	Case Western Reserve University
PU	Paderborn University
CALCE	Center for Advanced Life Cycle Engineerin
RUL	Remaining Useful Lifetime
C-MAPPS	Commercial Modular Aero-Propulsion System Simulation
EEG	Electroencephalogram
BCI	Brain-Computer Interface
HAR	Human Activity Recognition
SEED	Shanghai JiaoTong University Emotion EEG Dataset
DEAP	Database for Emotion Analysis using Physiological Signals
HHAR	Heterogeneity Human Activity Recognition
RKHS	Reproducing Kernel Hilbert Spaces
MMD	Maximum Mean Discrepancy
MK-MMD	Multi-kernel Maximum Mean Discrepancy

CORAL	Correlation Alignment
PAMAP	Physical Activity Monitoring for Aging People
CASIA	Institute of Automation, Chinese Academy of Science
OU-ISIR	Institute of Scientific and Industrial Research, Osaka University
MCI	Muscle-Computer Interface
EMG	Electromyograph
IMS	Intelligent Maintenance System

References

1. Javaid, M.; Haleem, A.; Rab, S.; Singh, R.P.; Suman, R. Sensors for daily life: A review. *Sens. Int.* **2021**, *2*, 100121. [[CrossRef](#)]
2. Heikenfeld, J.; Jajack, A.; Rogers, J.; Gutruf, P.; Tian, L.; Pan, T.; Li, R.; Khine, M.; Kim, J.; Wang, J. Wearable sensors: Modalities, challenges, and prospects. *Lab Chip* **2018**, *18*, 217–248. [[CrossRef](#)] [[PubMed](#)]
3. Arampatzis, T.; Lygeros, J.; Manesis, S. A survey of applications of wireless sensors and wireless sensor networks. In Proceedings of the International Symposium on, Mediterrean Conference on Control and Automation Intelligent Control, Limassol, Cyprus, 27–29 June 2005; pp. 719–724.
4. Fleming, W.J. New automotive sensors: A review. *Sens. J.* **2008**, *8*, 1900–1921. [[CrossRef](#)]
5. Kalsoom, T.; Ramzan, N.; Ahmed, S.; Ur-Rehman, M. Advances in sensor technologies in the era of smart factory and industry 4.0. *Sensors* **2020**, *20*, 6783. [[CrossRef](#)] [[PubMed](#)]
6. Angelopoulos, A.; Michailidis, E.T.; Nomikos, N.; Trakadas, P.; Hatziefremidis, A.; Voliotis, S.; Zahariadis, T. Tackling faults in the industry 4.0 era: A survey of machine-learning solutions and key aspects. *Sensors* **2019**, *20*, 109. [[CrossRef](#)] [[PubMed](#)]
7. Hasan, M.J.; Islam, M.M.; Kim, J.M. Bearing fault diagnosis using multidomain fusion-based vibration imaging and multitask learning. *Sensors* **2021**, *22*, 56. [[CrossRef](#)]
8. Fayyad, J.; Jaradat, M.A.; Gruyer, D.; Najjaran, H. Deep learning sensor fusion for autonomous vehicle perception and localization: A review. *Sensors* **2020**, *20*, 4220. [[CrossRef](#)]
9. Boukabache, H.; Escriba, C.; Fourniols, J.Y. Toward smart aerospace structures: Design of a piezoelectric sensor and its analog interface for flaw detection. *Sensors* **2014**, *14*, 20543–20561. [[CrossRef](#)]
10. Pytka, J.; Budzyński, P.; Łyszczczyk, T.; Józwick, J.; Michałowska, J.; Tofil, A.; Błażejczak, D.; Laskowski, J. Determining wheel forces and moments on aircraft landing gear with a dynamometer sensor. *Sensors* **2019**, *20*, 227. [[CrossRef](#)]
11. Kocić, J.; Jovičić, N.; Drndarević, V. Sensors and sensor fusion in autonomous vehicles. In Proceedings of the Telecommunications Forum, Belgrade, Serbia, 20–21 November 2018; pp. 420–425.
12. Varghese, J.Z.; Boone, R.G. Overview of autonomous vehicle sensors and systems. In Proceedings of the International Conference on Operations Excellence and Service Engineering, Orlando, FL, USA, 10–11 September 2015; pp. 178–191.
13. Van Brummelen, J.; O'Brien, M.; Gruyer, D.; Najjaran, H. Autonomous vehicle perception: The technology of today and tomorrow. *Transp. Res. Part C Emerg. Technol.* **2018**, *89*, 384–406. [[CrossRef](#)]
14. Voinea, G.D.; Butnariu, S.; Mogan, G. Measurement and geometric modelling of human spine posture for medical rehabilitation purposes using a wearable monitoring system based on inertial sensors. *Sensors* **2016**, *17*, 3. [[CrossRef](#)] [[PubMed](#)]
15. Arikumar, K.; Prathiba, S.B.; Alazab, M.; Gadekallu, T.R.; Pandya, S.; Khan, J.M.; Moorthy, R.S. FL-PMI: Federated learning-based person movement identification through wearable devices in smart healthcare systems. *Sensors* **2022**, *22*, 1377. [[CrossRef](#)] [[PubMed](#)]
16. Kwon, Y.H.; Shin, S.B.; Kim, S.D. Electroencephalography based fusion two-dimensional convolution neural networks model for emotion recognition system. *Sensors* **2018**, *18*, 1383. [[CrossRef](#)] [[PubMed](#)]
17. Leal-Junior, A.G.; Diaz, C.A.; Avellar, L.M.; Pontes, M.J.; Marques, C.; Frizzera, A. Polymer optical fiber sensors in healthcare applications: A comprehensive review. *Sensors* **2019**, *19*, 3156. [[CrossRef](#)]
18. Baskar, S.; Shakeel, P.M.; Kumar, R.; Burhanuddin, M.; Sampath, R. A dynamic and interoperable communication framework for controlling the operations of wearable sensors in smart healthcare applications. *Comput. Commun.* **2020**, *149*, 17–26. [[CrossRef](#)]
19. Ghaderpour, E.; Vujadinovic, T.; Hassan, Q.K. Application of the least-squares wavelet software in hydrology: Athabasca River basin. *J. Hydrol. Reg. Stud.* **2021**, *36*, 100847. [[CrossRef](#)]
20. Ghaderpour, E.; Pagiatakis, S.D.; Hassan, Q.K. A survey on change detection and time series analysis with applications. *Appl. Sci.* **2021**, *11*, 6141. [[CrossRef](#)]
21. Ghaderpour, E.; Pagiatakis, S.D. Least-squares wavelet analysis of unequally spaced and non-stationary time series and its applications. *Math. Geosci.* **2017**, *49*, 819–844. [[CrossRef](#)]
22. Krizhevsky, A.; Sutskever, I.; Hinton, G.E. Imagenet classification with deep convolutional neural networks. In Proceedings of the 25th International Conference on Neural Information Processing Systems, Lake Tahoe, NV, USA, 3–6 December 2012; Volume 1.
23. Wu, G.; Liu, Y.; Fang, L.; Chai, T. Revisiting light field rendering with deep anti-aliasing neural network. *arXiv* **2021**, arXiv:2104.06797.
24. Ismail Fawaz, H.; Forestier, G.; Weber, J.; Idoumghar, L.; Muller, P.A. Deep learning for time series classification: A review. *Data Min. Knowl. Discov.* **2019**, *33*, 917–963. [[CrossRef](#)]
25. Hochreiter, S.; Schmidhuber, J. Long short-term memory. *Neural Comput.* **1997**, *9*, 1735–1780. [[CrossRef](#)] [[PubMed](#)]

26. Li, M.; Zhang, M.; Luo, X.; Yang, J. Combined long short-term memory based network employing wavelet coefficients for MI-EEG recognition. In Proceedings of the International Conference on Mechatronics and Automation, Harbin, China, 7–10 August 2016; pp. 1971–1976.
27. Du, X.; Ma, C.; Zhang, G.; Li, J.; Lai, Y.K.; Zhao, G.; Deng, X.; Liu, Y.J.; Wang, H. An efficient LSTM network for emotion recognition from multichannel EEG signals. *Trans. Affect. Comput.* **2020**. [[CrossRef](#)]
28. Cho, K.; Van Merriënboer, B.; Gulcehre, C.; Bahdanau, D.; Bougares, F.; Schwenk, H.; Bengio, Y. Learning phrase representations using RNN encoder-decoder for statistical machine translation. *arXiv* **2014**, arXiv:1406.1078.
29. Yan, X.; Yang, J.; Song, L.; Liu, Y. PSA-GRU: Modeling person-social twin-attention based on GRU for pedestrian trajectory prediction. In Proceedings of the Chinese Control Conference, Shanghai, China, 26–28 July 2021; pp. 8151–8157.
30. Vaswani, A.; Shazeer, N.; Parmar, N.; Uszkoreit, J.; Jones, L.; Gomez, A.N.; Kaiser, Ł.; Polosukhin, I. Attention is all you need. In Proceedings of the 31st International Conference on Neural Information Processing Systems, Long Beach, CA, USA, 4–9 December 2017; Volume 30.
31. Kingma, D.P.; Ba, J. Adam: A method for stochastic optimization. *arXiv* **2014**, arXiv:1412.6980.
32. Ruder, S. An overview of gradient descent optimization algorithms. *arXiv* **2016**, arXiv:1609.04747.
33. Gonzalez-Jimenez, D.; Del-Olmo, J.; Poza, J.; Garramiola, F.; Madina, P. Data-driven fault diagnosis for electric drives: A review. *Sensors* **2021**, *21*, 4024. [[CrossRef](#)]
34. Torres, E.P.; Torres, E.A.; Hernández-Álvarez, M.; Yoo, S.G. EEG-based BCI emotion recognition: A survey. *Sensors* **2020**, *20*, 5083. [[CrossRef](#)]
35. Sousa Lima, W.; Souto, E.; El-Khatib, K.; Jalali, R.; Gama, J. Human activity recognition using inertial sensors in a smartphone: An overview. *Sensors* **2019**, *19*, 3213. [[CrossRef](#)]
36. Daumé, H., III. Frustratingly easy domain adaptation. *arXiv* **2009**, arXiv:0907.1815.
37. Zhao, J.; Deng, F.; He, H.; Chen, J. Local domain adaptation for cross-domain activity recognition. *Trans. Hum.-Mach. Syst.* **2021**, *5*, 12–21. [[CrossRef](#)]
38. Smith, W.A.; Randall, R.B. Rolling element bearing diagnostics using the Case Western Reserve University data: A benchmark study. *Mech. Syst. Signal Process.* **2015**, *64*, 100–131. [[CrossRef](#)]
39. Lessmeier, C.; Kimotho, J.K.; Zimmer, D.; Sextro, W. Condition monitoring of bearing damage in electromechanical drive systems by using motor current signals of electric motors: A benchmark data set for data-driven classification. In Proceedings of the PHM Society European Conference, Bilbao, Spain, 5–8 July 2016; Volume 3.
40. Lee, J.; Qiu, H.; Yu, G.; Lin, J. *Rexnord Technical Services: Bearing Data Set*; IMS, University of Cincinnati, NASA Ames Prognostics Data Repository: Moffett Field, CA, USA, 2007.
41. Xia, Y.; Shen, C.; Wang, D.; Shen, Y.; Huang, W.; Zhu, Z. Moment matching-based intraclass multisource domain adaptation network for bearing fault diagnosis. *Mech. Syst. Signal Process.* **2022**, *168*, 108697. [[CrossRef](#)]
42. Li, X.; Zhang, W.; Xu, N.X.; Ding, Q. Deep learning-based machinery fault diagnostics with domain adaptation across sensors at different places. *Trans. Ind. Electron.* **2019**, *67*, 6785–6794. [[CrossRef](#)]
43. Li, X.; Zhang, W.; Ding, Q.; Li, X. Diagnosing rotating machines with weakly supervised data using deep transfer learning. *Trans. Ind. Inform.* **2019**, *16*, 1688–1697. [[CrossRef](#)]
44. Zheng, H.; Wang, R.; Yang, Y.; Li, Y.; Xu, M. Intelligent fault identification based on multisource domain generalization towards actual diagnosis scenario. *Trans. Ind. Electron.* **2019**, *67*, 1293–1304. [[CrossRef](#)]
45. Zhu, J.; Chen, N.; Shen, C. A new multiple source domain adaptation fault diagnosis method between different rotating machines. *Trans. Ind. Inform.* **2020**, *17*, 4788–4797. [[CrossRef](#)]
46. Xiao, D.; Qin, C.; Yu, H.; Huang, Y.; Liu, C.; Zhang, J. Unsupervised machine fault diagnosis for noisy domain adaptation using marginal denoising autoencoder based on acoustic signals. *Measurement* **2021**, *176*, 109186. [[CrossRef](#)]
47. Xia, Y.; Shen, C.; Chen, Z.; Kong, L.; Huang, W.; Zhu, Z. Multisource domain transfer learning for bearing fault diagnosis. In Proceedings of the International Conference on Sensing, Measurement & Data Analytics in the Era of Artificial Intelligence, Nanjing, China, 21–23 October 2021; pp. 1–6.
48. Wang, Q.; Taal, C.; Fink, O. Integrating expert knowledge with domain adaptation for unsupervised fault diagnosis. *arXiv* **2021**, arXiv:2107.01849.
49. Liu, Z.; Lu, B.; Wei, H.; Chen, L.; Li, X.; Rättsch, M. Deep adversarial domain adaptation model for bearing fault diagnosis. *Trans. Syst. Man, Cybern. Syst.* **2019**, *51*, 4217–4226. [[CrossRef](#)]
50. Ragab, M.; Chen, Z.; Wu, M.; Li, H.; Kwok, C.K.; Yan, R.; Li, X. Adversarial multiple-target domain adaptation for fault classification. *Trans. Instrum. Meas.* **2020**, *70*, 1–11. [[CrossRef](#)]
51. Wang, Z.; Xia, H.; Zhu, S.; Peng, B.; Zhang, J.; Jiang, Y.; Annor-Nyarko, M. Cross-domain fault diagnosis of rotating machinery in nuclear power plant based on improved domain adaptation method. *J. Nucl. Sci. Technol.* **2022**, *59*, 67–77. [[CrossRef](#)]
52. Rezaeianjouybari, B.; Shang, Y. A novel deep multi-source domain adaptation framework for bearing fault diagnosis based on feature-level and task-specific distribution alignment. *Measurement* **2021**, *178*, 109359. [[CrossRef](#)]
53. Kim, M.; Ko, J.U.; Lee, J.; Youn, B.D.; Jung, J.H.; Sun, K.H. A domain adaptation with semantic clustering method for fault diagnosis of rotating machinery. *ISA Trans.* **2022**, *120*, 372–382. [[CrossRef](#)] [[PubMed](#)]
54. Li, X.; Zhang, W.; Ding, Q.; Sun, J.Q. Multi-layer domain adaptation method for rolling bearing fault diagnosis. *Signal Process.* **2019**, *157*, 180–197. [[CrossRef](#)]

55. Wei, D.; Han, T.; Chu, F.; Zuo, M.J. Weighted domain adaptation networks for machinery fault diagnosis. *Mech. Syst. Signal Process.* **2021**, *158*, 107744. [[CrossRef](#)]
56. Song, Y.; Li, Y.; Jia, L.; Qiu, M. Retraining strategy-based domain adaptation network for intelligent fault diagnosis. *Trans. Ind. Inform.* **2019**, *16*, 6163–6171. [[CrossRef](#)]
57. Yang, B.; Li, Q.; Chen, L.; Shen, C. Bearing fault diagnosis based on multilayer domain adaptation. *Shock Vib.* **2020**, *2020*, 8873960. [[CrossRef](#)]
58. Zhang, K.; Chen, J.; Zhang, T.; He, S.; Pan, T.; Zhou, Z. Intelligent fault diagnosis of mechanical equipment under varying working condition via iterative matching network augmented with selective Signal reuse strategy. *J. Manuf. Syst.* **2020**, *57*, 400–415. [[CrossRef](#)]
59. Xu, G.; Huang, C.; da Silva, D.S.; de Albuquerque, V.H.C. A compressed unsupervised deep domain adaptation model for efficient cross-domain fault diagnosis. *Trans. Ind. Inform.* **2022**. [[CrossRef](#)]
60. Wang, X.; He, H.; Li, L. A hierarchical deep domain adaptation approach for fault diagnosis of power plant thermal system. *Trans. Ind. Inform.* **2019**, *15*, 5139–5148. [[CrossRef](#)]
61. Wang, X.; Jin, M. Convolutional domain adaptation network for fault diagnosis of thermal system under different loading conditions. In Proceedings of the Chinese Control Conference, Online, 27–29 July 2020; pp. 4193–4197.
62. Pandhare, V.; Li, X.; Miller, M.; Jia, X.; Lee, J. Intelligent diagnostics for ball screw fault through indirect sensing using deep domain adaptation. *Trans. Instrum. Meas.* **2020**, *70*, 1–11. [[CrossRef](#)]
63. Azamfar, M.; Li, X.; Lee, J. Intelligent ball screw fault diagnosis using a deep domain adaptation methodology. *Mech. Mach. Theory* **2020**, *151*, 103932. [[CrossRef](#)]
64. Bole, B.; Kulkarni, C.S.; Daigle, M. Adaptation of an electrochemistry-based li-ion battery model to account for deterioration observed under randomized use. In Proceedings of the Annual Conference of the Prognostics and Health Management Society, Fort Worth, TX, USA, 29 September–2 October 2014; Volume 6.
65. Diao, W.; Saxena, S.; Pecht, M. Accelerated cycle life testing and capacity degradation modeling of LiCoO₂-graphite cells. *J. Power Sources* **2019**, *435*, 226830. [[CrossRef](#)]
66. Han, T.; Wang, Z.; Meng, H. End-to-end capacity estimation of Lithium-ion batteries with an enhanced long short-term memory network considering domain adaptation. *J. Power Sources* **2022**, *520*, 230823. [[CrossRef](#)]
67. Ye, Z.; Yu, J. State-of-health estimation for lithium-ion batteries using domain adversarial transfer learning. *Trans. Power Electron.* **2021**, *37*, 3528–3543. [[CrossRef](#)]
68. Saxena, A.; Goebel, K.; Simon, D.; Eklund, N. Damage propagation modeling for aircraft engine run-to-failure simulation. In Proceedings of the International Conference on Prognostics and Health Management, Denver, CO, USA, 6–9 October 2008; pp. 1–9.
69. Nectoux, P.; Gouriveau, R.; Medjaher, K.; Ramasso, E.; Chebel-Morello, B.; Zerhouni, N.; Varnier, C. PRONOSTIA: An experimental platform for bearings accelerated degradation tests. In Proceedings of the International Conference on Prognostics and Health Management, Beijing, China, 23–25 May 2012; pp. 1–8.
70. Zhuang, J.; Jia, M.; Ding, Y.; Ding, P. Temporal convolution-based transferable cross-domain adaptation approach for remaining useful life estimation under variable failure behaviors. *Reliab. Eng. Syst. Saf.* **2021**, *216*, 107946. [[CrossRef](#)]
71. Zhu, J.; Chen, N.; Shen, C. A new data-driven transferable remaining useful life prediction approach for bearing under different working conditions. *Mech. Syst. Signal Process.* **2020**, *139*, 106602. [[CrossRef](#)]
72. Wen, B.C.; Xiao, M.Q.; Wang, X.Q.; Zhao, X.; Li, J.F.; Chen, X. Data-driven remaining useful life prediction based on domain adaptation. *PeerJ Comput. Sci.* **2021**, *7*, e690. [[CrossRef](#)]
73. Fan, Y.; Nowaczyk, S.; Rögngvaldsson, T. Transfer learning for remaining useful life prediction based on consensus self-organizing models. *Reliab. Eng. Syst. Saf.* **2020**, *203*, 107098. [[CrossRef](#)]
74. Ragab, M.; Chen, Z.; Wu, M.; Kwok, C.K.; Li, X. Adversarial transfer learning for machine remaining useful life prediction. In Proceedings of the International Conference on Prognostics and Health Management, Detroit, MI, USA, 8–10 June 2020; pp. 1–7.
75. Ragab, M.; Chen, Z.; Wu, M.; Foo, C.S.; Kwok, C.K.; Yan, R.; Li, X. Contrastive adversarial domain adaptation for machine remaining useful life prediction. *Trans. Ind. Inform.* **2020**, *17*, 5239–5249. [[CrossRef](#)]
76. Fu, S.; Zhang, Y.; Lin, L.; Zhao, M.; Zhong, S.S. Deep residual LSTM with domain-invariance for remaining useful life prediction across domains. *Reliab. Eng. Syst. Saf.* **2021**, *216*, 108012. [[CrossRef](#)]
77. da Costa, P.R.d.O.; Akçay, A.; Zhang, Y.; Kaymak, U. Remaining useful lifetime prediction via deep domain adaptation. *Reliab. Eng. Syst. Saf.* **2020**, *195*, 106682. [[CrossRef](#)]
78. Zhuang, J.; Jia, M.; Ding, Y.; Ding, P. Remaining useful life estimation under variable failure behaviors via transferable metric learning. In Proceedings of the International Conference on Sensing, Measurement & Data Analytics in the era of Artificial Intelligence, Nanjing, China, 21–23 October 2021; pp. 1–6.
79. Ding, Y.; Jia, M.; Cao, Y. Remaining useful life estimation under multiple operating conditions via deep subdomain adaptation. *Trans. Instrum. Meas.* **2021**, *70*, 1–11. [[CrossRef](#)]
80. Liu, C.; Gryllias, K. Unsupervised domain adaptation based remaining useful life prediction of rolling element bearings. In Proceedings of the PHM Society European Conference, Virtual, 9–13 November 2020; Volume 5, p. 10.
81. Miao, M.; Yu, J. A deep domain adaptive network for remaining useful life prediction of machines under different working conditions and fault modes. *Trans. Instrum. Meas.* **2021**, *70*, 1–14. [[CrossRef](#)]

82. Li, Q.; Tang, B.; Deng, L.; Wu, Y.; Wang, Y. Deep balanced domain adaptation neural networks for fault diagnosis of planetary gearboxes with limited labeled data. *Measurement* **2020**, *156*, 107570. [[CrossRef](#)]
83. Zhang, M.; Lu, W.; Yang, J.; Wang, D.; Bin, L. Domain adaptation with multilayer adversarial learning for fault diagnosis of gearbox under multiple operating conditions. In Proceedings of the Prognostics and System Health Management Conference (PHM-Qingdao), Qingdao, China, 25–27 October 2019; pp. 1–6.
84. Qin, Y.; Yao, Q.; Wang, Y.; Mao, Y. Parameter sharing adversarial domain adaptation networks for fault transfer diagnosis of planetary gearboxes. *Mech. Syst. Signal Process.* **2021**, *160*, 107936. [[CrossRef](#)]
85. Zhao, B.; Zhang, X.; Zhan, Z.; Wu, Q. Deep multi-scale adversarial network with attention: A novel domain adaptation method for intelligent fault diagnosis. *J. Manuf. Syst.* **2021**, *59*, 565–576. [[CrossRef](#)]
86. Singh, J.; Azamfar, M.; Ainapure, A.; Lee, J. Deep learning-based cross-domain adaptation for gearbox fault diagnosis under variable speed conditions. *Meas. Sci. Technol.* **2020**, *31*, 055601. [[CrossRef](#)]
87. Zheng, W.; Lu, B. Personalizing EEG-based affective models with transfer learning. In Proceedings of the International Joint Conference on Artificial Intelligence, New York, NY, USA, 9–15 July 2016; pp. 2732–2738.
88. Koelstra, S.; Muhl, C.; Soleymani, M.; Lee, J.S.; Yazdani, A.; Ebrahimi, T.; Pun, T.; Nijholt, A.; Patras, I. Deap: A database for emotion analysis; Using physiological signals. *Trans. Affect. Comput.* **2011**, *3*, 18–31. [[CrossRef](#)]
89. Brunner, C.; Leeb, R.; Müller-Putz, G.; Schlögl, A.; Pfurtscheller, G. *BCI Competition 2008—Graz Data Set A*; Institute for Knowledge Discovery, Graz University of Technology: Graz, Austria 2008; Volume 16, pp. 1–6.
90. Leeb, R.; Brunner, C.; Müller-Putz, G.; Schlögl, A.; Pfurtscheller, G. *BCI Competition 2008—Graz Data Set B*; Graz University of Technology: Graz, Austria, 2008; pp. 1–6.
91. Jiménez-Guarneros, M.; Gómez-Gil, P. Custom domain adaptation: A new method for cross-subject, EEG-based cognitive load recognition. *Signal Process. Lett.* **2020**, *27*, 750–754. [[CrossRef](#)]
92. Raza, H.; Samothrakis, S. Bagging adversarial neural networks for domain adaptation in non-stationary eeg. In Proceedings of the International Joint Conference on Neural Networks, Budapest, Hungary, 14–19 July 2019; pp. 1–7.
93. Côté-Allard, U.; Gagnon-Turcotte, G.; Phinyomark, A.; Glette, K.; Scheme, E.J.; Laviolette, F.; Gosselin, B. Unsupervised domain adversarial self-calibration for electromyography-based gesture recognition. *IEEE Access* **2020**, *8*, 177941–177955. [[CrossRef](#)]
94. Ding, K.M.; Kimura, T.; Fukui, K.I.; Numao, M. Cross-phase Emotion Recognition using Multiple Source Domain Adaptation. In Proceedings of the BIOSIGNALS, Vienna, Austria, 11–13 February 2021; pp. 150–157.
95. Zhu, B.; Shoaran, M. Unsupervised domain adaptation for cross-subject few-shot neurological symptom detection. In Proceedings of the International Conference on Neural Engineering, Virtual, 4–6 May 2021; pp. 181–184.
96. He, Z.; Zhong, Y.; Pan, J. Joint temporal convolutional networks and adversarial discriminative domain adaptation for EEG-based cross-subject emotion recognition. In Proceedings of the International Conference on Acoustics, Speech and Signal Processing, Copenhagen, Denmark, 19–20 July 2022; pp. 3214–3218.
97. Ahmed, M.Z.I.; Sinha, N.; Phadikar, S.; Ghaderpour, E. Automated feature extraction on AsMap for emotion classification using eeg. *Sensors* **2022**, *22*, 2346. [[CrossRef](#)] [[PubMed](#)]
98. Phadikar, S.; Sinha, N.; Ghosh, R.; Ghaderpour, E. Automatic muscle artifacts identification and removal from single-channel eeg using wavelet transform with meta-heuristically optimized non-local means filter. *Sensors* **2022**, *22*, 2948. [[CrossRef](#)] [[PubMed](#)]
99. Szttyler, T.; Stuckenschmidt, H. On-body localization of wearable devices: An investigation of position-aware activity recognition. In Proceedings of the International Conference on Pervasive Computing and Communications, Sydney, Australia, 14–18 March 2016; pp. 1–9.
100. Roggen, D.; Calatroni, A.; Rossi, M.; Holleczeck, T.; Förster, K.; Tröster, G.; Lukowicz, P.; Bannach, D.; Pirkl, G.; Ferscha, A.; et al. Collecting complex activity datasets in highly rich networked sensor environments. In Proceedings of the International Conference on Networked Sensing Systems, Kassel, Germany, 15–18 June 2010; pp. 233–240.
101. Stisen, A.; Blunck, H.; Bhattacharya, S.; Prentow, T.S.; Kjærgaard, M.B.; Dey, A.; Sonne, T.; Jensen, M.M. Smart devices are different: Assessing and mitigating mobile sensing heterogeneities for activity recognition. In Proceedings of the Conference on Embedded Networked Sensor Systems, Seoul, Korea, 1–4 November 2015; pp. 127–140.
102. Reiss, A.; Stricker, D. Introducing a new benchmarked dataset for activity monitoring. In Proceedings of the International Symposium on Wearable Computers, Newcastle Upon Tyne, UK, 18–22 June 2012; pp. 108–109.
103. Hussein, A.; Hajj, H. Domain adaptation with representation learning and nonlinear relation for time series. *Trans. Internet Things* **2022**, *3*, 1–26. [[CrossRef](#)]
104. Sanabria, A.R.; Zambonelli, F.; Dobson, S.; Ye, J. ContrasGAN: Unsupervised domain adaptation in Human Activity Recognition via adversarial and contrastive learning. *Pervasive Mob. Comput.* **2021**, *78*, 101477. [[CrossRef](#)]
105. Zhang, H.; Zhou, Z.; Gong, W. Wi-adaptor: Fine-grained domain adaptation in wifi-based activity recognition. In Proceedings of the Global Communications Conference, Madrid, Spain, 7–11 December 2021; pp. 1–6.
106. Wang, F.; Liu, J.; Gong, W. WiCAR: WiFi-based in-car activity recognition with multi-adversarial domain adaptation. In Proceedings of the International Symposium on Quality of Service, Phoenix, AZ, USA, 24–25 June 2019; pp. 1–10.
107. Akbari, A.; Jafari, R. Transferring activity recognition models for new wearable sensors with deep generative domain adaptation. In Proceedings of the International Conference on Information Processing in Sensor Networks, Montreal, QC, Canada, 15–18 April 2019; pp. 85–96.

108. Chakma, A.; Faridee, A.Z.M.; Khan, M.A.A.H.; Roy, N. Activity recognition in wearables using adversarial multi-source domain adaptation. *Smart Health* **2021**, *19*, 100174. [[CrossRef](#)]
109. Sanabria, A.R.; Zambonelli, F.; Ye, J. Unsupervised domain adaptation in activity recognition: A GAN-based approach. *IEEE Access* **2021**, *9*, 19421–19438. [[CrossRef](#)]
110. Khan, M.A.A.H.; Roy, N.; Misra, A. Scaling human activity recognition via deep learning-based domain adaptation. In Proceedings of the International Conference on Pervasive Computing and Communications, Athens, Greece, 19–23 March 2018; pp. 1–9.
111. Du, Y.; Jin, W.; Wei, W.; Hu, Y.; Geng, W. Surface EMG-based inter-session gesture recognition enhanced by deep domain adaptation. *Sensors* **2017**, *17*, 458. [[CrossRef](#)]
112. Atzori, M.; Gijssberts, A.; Castellini, C.; Caputo, B.; Hager, A.G.M.; Elsig, S.; Giatsidis, G.; Bassetto, F.; Müller, H. Electromyography data for non-invasive naturally-controlled robotic hand prostheses. *Sci. Data* **2014**, *1*, 1–13. [[CrossRef](#)]
113. Amma, C.; Krings, T.; Böer, J.; Schultz, T. Advancing muscle-computer interfaces with high-density electromyography. In Proceedings of the Conference on Human Factors in Computing Systems, Seoul, Korea, 18–23 April 2015; pp. 929–938.
114. Chan, P.P.; Li, Q.; Fang, Y.; Xu, L.; Li, K.; Liu, H.; Yeung, D.S. Unsupervised domain adaptation for gesture identification against electrode shift. *Trans. Hum.-Mach. Syst.* **2022**. [[CrossRef](#)]
115. Ketykó, I.; Kovács, F.; Varga, K.Z. Domain adaptation for semg-based gesture recognition with recurrent neural networks. In Proceedings of the International Joint Conference on Neural Networks, Budapest, Hungary, 14–19 July 2019; pp. 1–7.
116. Liu, J.; Sheng, X.; Zhang, D.; He, J.; Zhu, X. Reduced daily recalibration of myoelectric prosthesis classifiers based on domain adaptation. *J. Biomed. Health Inform.* **2014**, *20*, 166–176. [[CrossRef](#)] [[PubMed](#)]
117. Zakia, U.; Menon, C. Force myography-based human robot interactions via deep domain adaptation and generalization. *Sensors* **2021**, *22*, 211. [[CrossRef](#)] [[PubMed](#)]
118. Bachlin, M.; Plotnik, M.; Roggen, D.; Maidan, I.; Hausdorff, J.M.; Giladi, N.; Troster, G. Wearable assistant for Parkinson’s disease patients with the freezing of gait symptom. *Trans. Inf. Technol. Biomed.* **2009**, *14*, 436–446. [[CrossRef](#)] [[PubMed](#)]
119. Iwama, H.; Okumura, M.; Makihara, Y.; Yagi, Y. The OU-ISIR gait database comprising the large population dataset and performance evaluation of gait recognition. *Trans. Inf. Forensics Secur.* **2012**, *7*, 1511–1521. [[CrossRef](#)]
120. Yu, S.; Tan, D.; Tan, T. A framework for evaluating the effect of view angle, clothing and carrying condition on gait recognition. In Proceedings of the International Conference on Pattern Recognition, Hong Kong, China, 20–24 August 2006; Volume 4, pp. 441–444.
121. Tan, D.; Huang, K.; Yu, S.; Tan, T. Efficient night gait recognition based on template matching. In Proceedings of the International Conference on Pattern Recognition, Hong Kong, China, 20–24 August 2006; Volume 3, pp. 1000–1003.
122. Yang, Y.; Yang, X.; Sakamoto, T.; Fioranelli, F.; Li, B.; Lang, Y. Unsupervised domain adaptation for disguised-gait-based person identification on micro-doppler signatures. *Trans. Circuits Syst. Video Technol.* **2022**. [[CrossRef](#)]
123. Zheng, J.; Liu, X.; Yan, C.; Zhang, J.; Liu, W.; Zhang, X.; Mei, T. Trand: Transferable neighborhood discovery for unsupervised cross-domain gait recognition. In Proceedings of the International Symposium on Circuits and Systems, Daegu, Korea, 22–28 May 2021; pp. 1–5.
124. Mu, F.; Gu, X.; Guo, Y.; Lo, B. Unsupervised domain adaptation for position-independent imu based gait analysis. In Proceedings of the SENSORS, Rotterdam, The Netherlands, 25–28 October 2020; pp. 1–4.
125. Torvi, V.G.; Bhattacharya, A.; Chakraborty, S. Deep domain adaptation to predict freezing of gait in patients with Parkinson’s disease. In Proceedings of the International Conference on Machine Learning and Applications, Orlando, FL, USA, 17–20 December 2018; pp. 1001–1006.
126. Choi, W.; Yang, W.; Na, J.; Park, J.; Lee, G.; Nam, W. Unsupervised gait phase estimation with domain-adversarial neural network and adaptive window. *J. Biomed. Health Inform.* **2021**, *26*, 3373–3384. [[CrossRef](#)]
127. Guo, Y.; Gu, X.; Yang, G.Z. MCDGD: Multi-source unsupervised domain adaptation for abnormal human gait detection. *J. Biomed. Health Inform.* **2021**, *25*, 4017–4028. [[CrossRef](#)]
128. Morales-Espejel, G.E.; Gabelli, A. Rolling bearing performance rating parameters: Review and engineering assessment. *Proc. Inst. Mech. Eng. Part C* **2020**, *234*, 3064–3077. [[CrossRef](#)]
129. Canizo, M.; Triguero, I.; Conde, A.; Onieva, E. Multi-head CNN–RNN for multi-time series anomaly detection: An industrial case study. *Neurocomputing* **2019**, *363*, 246–260. [[CrossRef](#)]
130. Martin, R.; Lazakis, I.; Barbouchi, S.; Johanning, L. Sensitivity analysis of offshore wind farm operation and maintenance cost and availability. *Renew. Energy* **2016**, *85*, 1226–1236. [[CrossRef](#)]
131. Shao, S.; McAleer, S.; Yan, R.; Baldi, P. Highly accurate machine fault diagnosis using deep transfer learning. *IEEE Trans. Ind. Inform.* **2018**, *15*, 2446–2455. [[CrossRef](#)]
132. Kurşun, B.; Ökten, K. Thermodynamic analysis of a Rankine cycle coupled with a concentrated photovoltaic thermal system for hydrogen production by a proton exchange membrane electrolyzer plant. *Int. J. Hydrog. Energy* **2019**, *44*, 22863–22875. [[CrossRef](#)]
133. Panwar, N.; Kaushik, S.; Kothari, S. Role of renewable energy sources in environmental protection: A review. *Renew. Sustain. Energy Rev.* **2011**, *15*, 1513–1524. [[CrossRef](#)]
134. Vashisht, R.K.; Peng, Q. Online chatter detection for milling operations using LSTM neural networks assisted by motor current signals of ball screw drives. *J. Manuf. Sci. Eng.* **2021**, *143*, 011008. [[CrossRef](#)]
135. Manthiram, A. A reflection on lithium-ion battery cathode chemistry. *Nat. Commun.* **2020**, *11*, 1550. [[CrossRef](#)] [[PubMed](#)]

136. Harper, G.; Somerville, R.; Kendrick, E.; Driscoll, L.; Slater, P.; Stolkin, R.; Walton, A.; Christensen, P.; Heidrich, O.; Lambert, S.; et al. Recycling lithium-ion batteries from electric vehicles. *Nature* **2019**, *575*, 75–86. [[CrossRef](#)]
137. Hannan, M.A.; Hoque, M.M.; Hussain, A.; Yusof, Y.; Ker, P.J. State-of-the-art and energy management system of lithium-ion batteries in electric vehicle applications: Issues and recommendations. *IEEE Access* **2018**, *6*, 19362–19378. [[CrossRef](#)]
138. Gandoman, F.H.; Jaguemont, J.; Goutam, S.; Gopalakrishnan, R.; Firouz, Y.; Kalogiannis, T.; Omar, N.; Van Mierlo, J. Concept of reliability and safety assessment of lithium-ion batteries in electric vehicles: Basics, progress, and challenges. *Appl. Energy* **2019**, *251*, 113343. [[CrossRef](#)]
139. Chandran, V.; Patil, C.K.; Karthick, A.; Ganeshaperumal, D.; Rahim, R.; Ghosh, A. State of charge estimation of lithium-ion battery for electric vehicles using machine learning algorithms. *World Electr. Veh. J.* **2021**, *12*, 38. [[CrossRef](#)]
140. Jayasinghe, L.; Samarasinghe, T.; Yuenv, C.; Low, J.C.N.; Ge, S.S. Temporal convolutional memory networks for remaining useful life estimation of industrial machinery. In Proceedings of the International Conference on Industrial Technology, Melbourne, Australia, 13–15 February 2019; pp. 915–920.
141. Yoo, Y.; Baek, J.G. A novel image feature for the remaining useful lifetime prediction of bearings based on continuous wavelet transform and convolutional neural network. *Appl. Sci.* **2018**, *8*, 1102. [[CrossRef](#)]
142. Salameh, J.P.; Cauet, S.; Etien, E.; Sakout, A.; Rambault, L. Gearbox condition monitoring in wind turbines: A review. *Mech. Syst. Signal Process.* **2018**, *111*, 251–264. [[CrossRef](#)]
143. Bhardwaj, U.; Teixeira, A.; Soares, C.G. Reliability prediction of an offshore wind turbine gearbox. *Renew. Energy* **2019**, *141*, 693–706. [[CrossRef](#)]
144. Elasha, F.; Greaves, M.; Mba, D. Planetary bearing defect detection in a commercial helicopter main gearbox with vibration and acoustic emission. *Struct. Health Monit.* **2018**, *17*, 1192–1212. [[CrossRef](#)]
145. Parvaresh, A.; Mardani, M. Model predictive control of a hydraulic actuator in torque applying system of a mechanically closed-loop test rig for the helicopter gearbox. *Aviation* **2019**, *23*, 143–153. [[CrossRef](#)]
146. Somani, N.; Zhen, L.L.; Lakshminarayanan, S.; Hettiarachchi, R.; Wee-Ching, P.; Lee, G.S.G.; Campolo, D. Simplified robot programming framework for a gearbox assembly application. In Proceedings of the International Conference on Social Robotics, Singapore, 10–13 November 2021; pp. 181–190.
147. Miklos, I.Z.; Miklos, C.C.; Alic, C.I. Dynamic simulation of a planetary gearbox with double satellite. In Proceedings of the International Conference on Machine and Industrial Design in Mechanical Engineering, Novi Sad, Serbia, 10–12 June 2021; pp. 523–530.
148. Hramov, A.E.; Maksimenko, V.A.; Pisarchik, A.N. Physical principles of brain–computer interfaces and their applications for rehabilitation, robotics and control of human brain states. *Phys. Rep.* **2021**, *918*, 1–133. [[CrossRef](#)]
149. Lotte, F.; Bougrain, L.; Cichocki, A.; Clerc, M.; Congedo, M.; Rakotomamonjy, A.; Yger, F. A review of classification algorithms for EEG-based brain–computer interfaces: A 10 year update. *J. Neural Eng.* **2018**, *15*, 031005. [[CrossRef](#)]
150. Phan, H.; Andreotti, F.; Cooray, N.; Chén, O.Y.; De Vos, M. DNN filter bank improves 1-max pooling CNN for single-channel EEG automatic sleep stage classification. In Proceedings of the International Conference of the Engineering in Medicine and Biology Society, Honolulu, HI, USA, 17–21 July 2018; pp. 453–456.
151. Tangermann, M.; Müller, K.R.; Aertsen, A.; Birbaumer, N.; Braun, C.; Brunner, C.; Leeb, R.; Mehring, C.; Miller, K.J.; Mueller-Putz, G.; et al. Review of the BCI competition IV. *Front. Neurosci.* **2012**, *6*, 55. [[CrossRef](#)]
152. Jobanputra, C.; Bavishi, J.; Doshi, N. Human activity recognition: A survey. *Procedia Comput. Sci.* **2019**, *155*, 698–703. [[CrossRef](#)]
153. Vrigkas, M.; Nikou, C.; Kakadiaris, I.A. A review of human activity recognition methods. *Front. Robot. AI* **2015**, *2*, 28. [[CrossRef](#)]
154. Hussein, D.; Bhat, G.; Doppa, J.R. Adaptive energy management for self-sustainable wearables in mobile health. In Proceedings of the AAAI Conference on Artificial Intelligence, Virtual, 22 February–1 March 2022.
155. Brock, A.; Donahue, J.; Simonyan, K. Large scale GAN training for high fidelity natural image synthesis. *arXiv* **2018**, arXiv:1809.11096.
156. Li, Y.; Wang, N.; Shi, J.; Liu, J.; Hou, X. Revisiting batch normalization for practical domain adaptation. *arXiv* **2016**, arXiv:1603.04779.
157. Gholami, B.; Sahu, P.; Rudovic, O.; Bousmalis, K.; Pavlovic, V. Unsupervised multi-target domain adaptation: An information theoretic approach. *Trans. Image Process.* **2020**, *29*, 3993–4002. [[CrossRef](#)] [[PubMed](#)]
158. Hinton, G.; Vinyals, O.; Dean, J. Distilling the knowledge in a neural network. *arXiv* **2015**, arXiv:1503.02531.
159. Chen, T.; Kornblith, S.; Norouzi, M.; Hinton, G. A simple framework for contrastive learning of visual representations. In Proceedings of the International Conference on Machine Learning, Vienna, Austria, 12–18 July 2020; pp. 1597–1607.