Information Fusion 118 (2025) 102982

Contents lists available at ScienceDirect

Information Fusion

journal homepage: www.elsevier.com/locate/inffus

Application of transfer learning for biomedical signals: A comprehensive review of the last decade (2014–2024)

Mahboobeh Jafari [a,*], Xiaohui Tao [a], Prabal Barua [a,b,c,d,e,f,g], Ru-San Tan [h,i], U.Rajendra Acharya [a]

[a] School of Mathematics, Physics and Computing, University of Southern Queensland, Springfield, QLD 4300, Australia
[b] School of Science and Technology, Kumamoto University, Japan
[c] Faculty of Engineering and Information Technology, University of Technology Sydney, Sydney, NSW 2007, Australia
[d] Australian International Institute of Higher Education, Sydney, NSW 2000, Australia
[e] Cogninet AI, Sydney, NSW 2010, Australia
[f] School of Science & Technology, University of New England, Australia
[g] School of Computing, SRM Institute of Science and Technology, India
[h] National Heart Centre Singapore, Singapore
[i] Duke-NUS Medical School, Singapore

ARTICLE INFO

Keywords:
Transfer learning
Pre-trained
Artificial intelligence
Biomedical signals
Machine learning
Deep learning
Disease diagnosis
BCIs

ABSTRACT

Precise and timely disease diagnosis is essential for making effective treatment decisions and halting disease progression. Biomedical signals offer the potential for non-invasive diagnosis of diverse conditions, enhancing the ability to predict clinical outcomes and plan treatments more effectively. These signals have garnered significant attention, particularly in +conjunction with artificial intelligence (AI)-powered models, such as conventional machine learning (ML) and deep learning (DL), demonstrating promising outcomes. However, DL models, which have become the de facto standard in medical data analysis, encounter challenges such as inadequate data availability, improper distribution, and storage limitations. To mitigate these issues, transfer learning (TL) has been employed to transfer knowledge from one domain to a related domain, enabling models to be fine-tuned with small-scale data while ensuring adaptability across diverse contexts, including variations in subjects, datasets, and sessions. This review presents a detailed and systematic overview of studies from the current decade that have employed TL models for healthcare-related applications using biomedical signals. In the introduction section, we explain the importance of employing TL techniques on biomedical signals in various domains, including disease diagnosis and prediction, and brain-computer interfaces (BCIs). The following section presents TL strategies. Another section is dedicated to searching and selection of articles based on the PRISMA method from reference databases including IEEE, Scopus, Web of Science, and PubMed. In this review, we examined 239 Q1 articles. Review articles published using TL techniques with biomedical signals are discussed in a separate section. In this review, we have studied the papers that have utilized TL techniques with various biosignals for various applications. Following this, we discuss the key challenges and future directions for the field based on the reviewed articles and conclude with a summary of key findings. Based on our study, EEG signals were the most frequently utilized in TL methods, particularly in the context of Brain-Computer Interface (BCI) applications, followed by applications in epilepsy detection. Additionally, domain adaptation methods are widely used in biomedical signals to address variations in data distribution caused by differences in subjects, devices, datasets, and recording conditions. These methods aim to align source and target domains, enabling models to generalize effectively across diverse datasets. This study provides a comprehensive review of current TL methods, offering useful insights for choosing the most suitable TL techniques for specific applications. It aims to deal with problems like data scarcity, domain mismatches, real-time issues, and hardware resource constraints in real-world scenarios.

* Corresponding author.
E-mail address: Mahboobeh.jafari@unisq.edu.au (M. Jafari).

https://doi.org/10.1016/j.inffus.2025.102982
Received 8 January 2025; Received in revised form 24 January 2025; Accepted 25 January 2025
Available online 30 January 2025

1. Introduction

Analysis of biomedical signals is widely exploited in healthcare applications for diagnosis, prediction, and therapeutic monitoring of diverse medical conditions [1,2]. These signals are captured using different sensors that are either implanted or positioned on specific regions of the skin surface [3], e.g., electromyography (EMG), which records surface skeletal muscle electrical activity; electrocardiography (ECG), surface heart muscle electrical activity, which reflects heart rate and rhythm; electroencephalogram (EEG), surface electrical activity of brain activity via scalp electrodes; photoplethysmography (PPG), blood volume within tissue vascular beds, which reflects oxygenation as well as pulsatile blood flow; and electrooculogram (EOG), corneal-retinal electrical potential. Extracting relevant features from these signals is obligatory for medical diagnostic or prognostic interpretation [4,5].

The interpretation of biomedical signals for healthcare presents several challenges. While some techniques like PPG is widely accessible, signal acquisition in some cases such as EEG and EMG require specialized skills and can be expensive. Expert analysis of signals is manually intensive and liable to fatigue, especially if the data volume is large, such as in continuous data recorded by wearable devices [6,7]. Moreover, altered signal waveform due to artifact and noise can compromise accurate interpretation [8,9]. Finally, even among experts, manual interpretation is subject to intra- and inter- differences among operators [10]. To tackle these challenges, many researchers have developed artificial intelligence (AI)-enabled computer-aided diagnosis (CAD) systems that can efficiently learn expert-labeled biomedical signals and are capable of automated classification for clinical diagnostic and predictive applications [11]. For instance, in the neuroscience field, EEG-based CAD systems have been applied to diverse applications, including emotion recognition [12,13,14]; detection of disorders like epilepsy [15–17], autism spectrum disorders (ASD) [18–20], attention deficit hyperactivity disorder (ADHD) [21–23], and schizophrenia (SZ) [24–26]; and brain-computer interfaces (BCIs) for physical rehabilitation [27–29].

Most AI CAD systems are based on either deep learning (DL) or traditional machine learning (ML). Compared with DL, ML offers explainability (as the feature engineering is designed with expert inputs), is better able to train using small data volumes, has lower risk of overfitting, and requires less computational power [30,31]. However, ML-based CAD systems may be inefficient in real-world scenarios and perform poorly with large datasets. Using deep architectures, DL models can effectively process large signal data volumes to automatically extract intricate, high-level features from raw signals, capturing patterns that might be overlooked by humans [32,33], obviating the need for signal preprocessing and prior design-intensive handcrafted feature engineering [34–37]. DL is widely adopted due to high performance but requires large training data volume, consistent training and test data distributions, and high time and computational costs. DL performs best when training with high-volume labeled training datasets [38]; and also, when the feature space and distribution of the training and test data are the same, which necessitates model reconstruction using updated data should the distribution become altered [36]. Whether collecting medical training data from scratch or recollecting training data (and then rebuilding models) are tasks that are neither cheap nor practical for many medical conditions [37]. By leveraging large amounts of unlabeled data to enhance learning accuracy with limited amount of labeled data, semi-supervised learning partially mitigates the data scarcity problem by minimizing the requirement for large amounts of labeled training data [38]. Nevertheless, collecting unlabeled samples can also be challenging, often resulting in suboptimal models. Alternative solutions such as data augmentation, data synthesis, cloud computing, and distributed learning have been proposed to address these limitations [36] but these approaches frequently come with trade-offs regarding computational demands, efficiency, and security.

Transfer learning (TL) has gained attention as a promising solution to DL's present challenges. Utilizing knowledge gained from source tasks in different domains, TL eliminates the need to learn from scratch with large training dataset [39]. This approach addresses the critical issue of insufficient labeled medical training data [40,41]. As TL has been pre-trained, time and computational resources required for model training are considerably decreased. When data comes from dissimilar sources or conditions (e.g., different people, devices, or environments), it often varies in how it is distributed, leading to mismatches which can lessen the accuracy and effectiveness of ML models, especially when the testing data differs from the training data. TL addresses this challenge by transferring knowledge from one domain (e.g., a dataset of ECG signals) to another or even combining knowledge from multiple domains. This allows models to adapt better to new or varied data, improving their performance and ensuring they work effectively across diverse sources or distributions [39].

This work aims to systematically review TL and its implementation in healthcare applications based on diverse biomedical signals, which may benefit researchers in this field. The article is structured as follows: Section 2 explains the different categories of TL and the principles underlying each; Section 3 explains the review method and compares our work against existing literature; Section 4 detains to results, stratified by signal types; Section 5 discusses challenges and future works; and Section 6 outlines our conclusions.

2. Transfer learning

DL models assume that there is a similarity in feature spaces and distributions between the training data (source domain) and the test data (target domain). When there is a discrepancy in the feature space or distribution of test data, existing prediction or detection models become ineffective. This requires rebuilding and retraining models with newly acquired training data, which can be costly, and impractical. Establishing such models for target domains with limited labeled data is exceedingly challenging through supervised learning alone: it is often difficult to gather enough labeled data, making it impossible to build new models from scratch through data collection. With TL, which transfers knowledge from another domain that possesses abundant labeled data, we can circumvent the limitation to facilitate development of learning-based models for the target domain.

In developing DL models for healthcare applications, there is primarily a lack of domain-specific labeled medical data, which require specialized laborious and costly labeling by experts. In contrast, the use of natural images for model training is not limited by scarcity, access, and ethical constraints [42]. TL, which uses networks pre-trained on natural images, has unsurprisingly burgeoned in AI research. The pre-trained models like LeNet [43], AlexNet [44], VGGNet [45], ResNet [46], GoogLeNet [47], DenseNet [48], XceptionNet [49], and SqueezeNet [50] have been fully trained on ImageNet [51]. By retraining certain aspects of these networks to adapt them to new domains, TL provides faster convergence and satisfactory results. While TL has proven successful in numerous applications, the distinct characteristics of medical data raise concerns and uncertainties.

2.1. Transfer learning advantages

TL confers several advantages. So long as there are similarities between the source and target domains, the knowledge gained by the model trained on the source domain, which typically contains abundant labeled data, can be leveraged in the target domain, even though the latter may have limited or no labeled data (Fig. 1) [52–54]. This approach reduces dependency on large, annotated datasets, minimizing the need for extensive medical data collection and labeling, which can be expensive and difficult to obtain [55]. Besides requiring less training data, TL mitigates issues like overfitting as well as enables the development of deeper networks with more parameters, which may improve target models. On limited target domain data, fine-tuned pre-trained models often extract significant features that result in superior

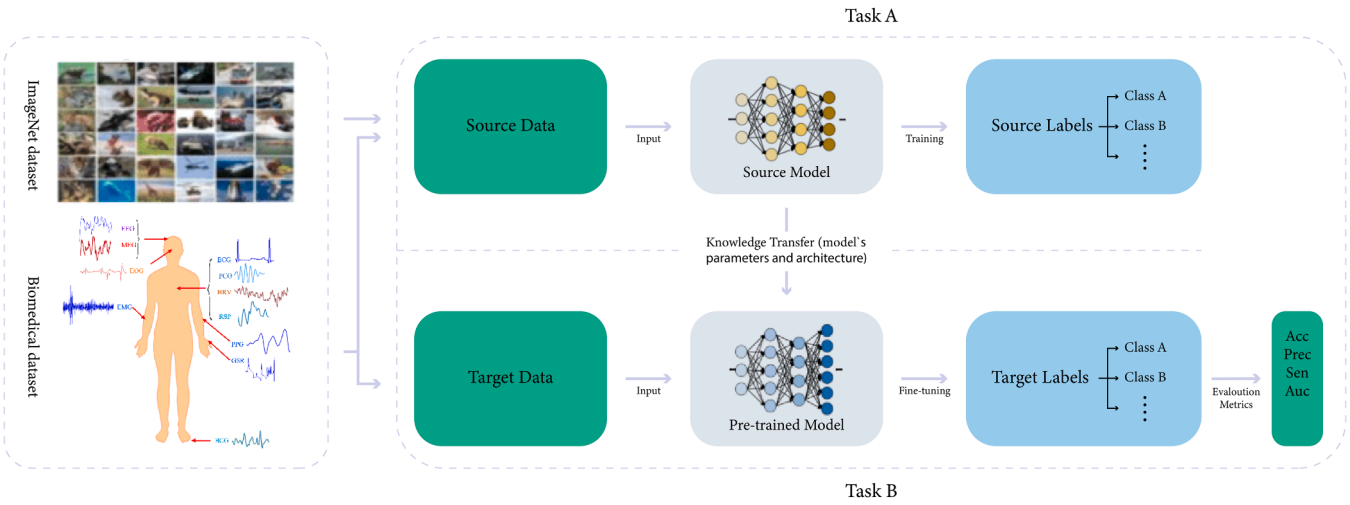

Fig. 1. Workflow of transfer learning using biomedical signals.

performance compared with training models from scratch [56,57].

Pre-trained models can optimize their transferred parameters to the target domain using small amounts of data, thereby accelerating convergence of target models, and reducing training time and computational costs. By enabling efficient deployment on resource-constrained hardware, TL supports practical applications requiring streamlined computations [58], e.g., continuous wearable monitors, which possess limited processing power and memory. Finally, TL enhances the generalizability of models. Biomedical signals often show significant variability across individuals and sessions, which can degrade model performance and require re-training for different users and sessions. TL helps address this issue by extracting transferable features that can be applied across different tasks and subjects. Methods like domain adaptation further enhance generalizability, allowing models to perform well across diverse scenarios without needing to retrain for every new user or session. By creating reusable models, TL reduces training demands and improves reliability, turning it into an affordable and efficient solution for diverse applications [59].

2.2. Types of transfer learning

There is no consensus on a unified or standardized approach for classifying the different types of TL. To facilitate understanding of TL from various perspectives, we have grouped TL into categories of TL according to the label, feature, and learning style in this work.

2.2.1. Sort by label types

Depending on whether the labeled data is available in the source or target domain, TL can be divided into inductive transductive, and unsupervised TL (Fig. 2) [53,60].

In inductive TL, there is a difference in tasks between the source and target domains, but they remain related, and the target domain has limited labeled data. Two possible scenarios exist. If the source domain incorporates substantial labeled data, it is akin to multitask learning albeit with a difference: whereas multitask learning simultaneously learns both source and target tasks, inductive TL improve the target task performance using knowledge gained from the source task [53]. If the source domain contains little or no labeled data, it becomes like self-taught learning [20,61].

In transductive TL, the tasks in the source domain and target domain are the same; and the target domain lacks labeled data. The latter makes it well-suited for medical practices, as labeled target domain data are often unavailable. Transductive TL can be subclassified according to the level of resemblance in feature spaces across the source and target domains. Where both the domains and tasks in the source and target are identical, methods addressing sample selection bias [62] or covariate shift [63] are used. If the domains differ but the tasks are the same, the focus shifts to situations where the marginal probability distributions of

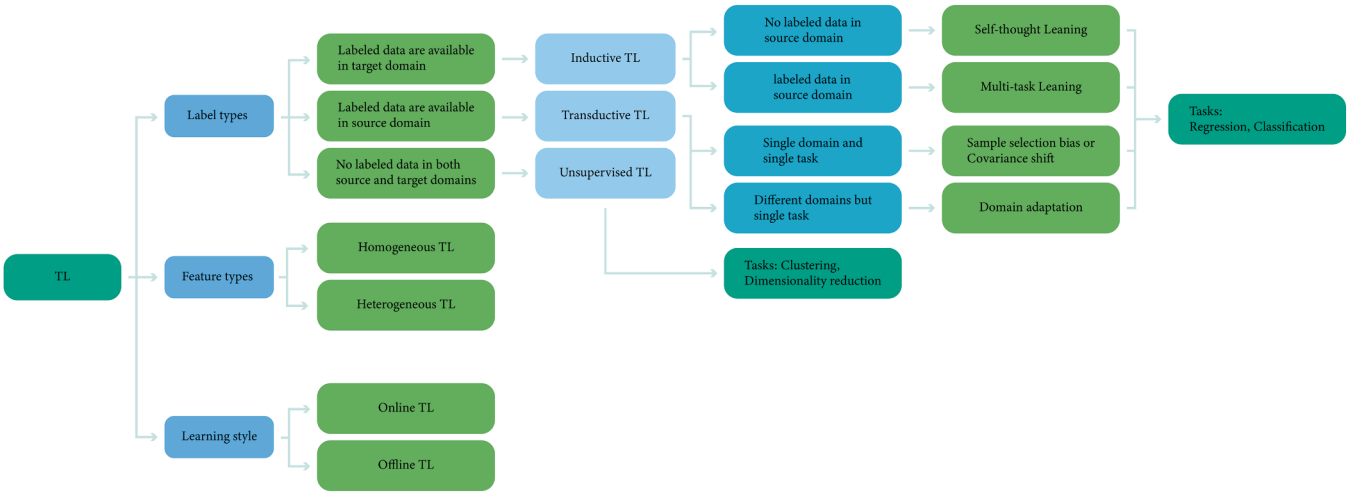

Fig. 2. Overview of TL strategies.

the input data vary [53,64]. The latter case is relevant to a heterogeneous TL method, domain adaptation, which is used for knowledge transfer in various cross-subject, cross-session, and cross-dataset applications, e.g., EEG-based emotion recognition [65,66] and BCI [67], and ECG-based arrhythmia classification [68]. Domain adaptation [69] involves sufficient training on source data, followed by adjustments based on distribution of target data.

In unsupervised TL, the tasks in the source domain and target domain differ; and both the source domain and target domain contain unlabeled data. It leverages correlation between the source and target domains to facilitate unsupervised learning tasks in the target domain, including clustering and dimensionality reduction [70]. Unavailability of large, labeled datasets and multiple co-occurring diseases, which complicate the labeling process, are issues that confound supervised learning. These challenges are less impactful for unsupervised approaches, as they can learn effectively from substantial amounts of unlabeled data via pre-training. If sufficient features are extracted during pre-training, models can achieve high performance in downstream tasks despite limited data. In [71], unsupervised TL models applied to analyzes of EEG and ECG signals yielded promising results. Unsupervised learning methods rely on generative models. e.g., variational autoencoders (VAE), clustering, and generative adversarial networks (GANs).

2.2.2. Sort by feature types

Depending on feature differences between the source domain and target domain, TL can be divided into homogeneous TL and heterogeneous TL (Fig. 2) [36,53]. In homogeneous TL, the source domain and target domain share the same feature space in terms of semantics and dimensions and is generally easier to implement. In heterogeneous TL, feature sets in the source domain and target domain differ in semantics and dimensions [52]. The latter more resembles real-world scenarios, in which highly heterogeneous biomedical signals result in variations in feature spaces and distributions between the source and target domains. Heterogeneous TL plays a crucial role in addressing these domain differences and shifts, enabling algorithms to be employed across a range of healthcare applications without the need for advanced techniques to align the distributions of training and test data.

2.2.3. Sort by learning style

TL can be conducted offline or online (Fig. 2). In offline TL, the source and target domains are static; a single transfer of knowledge occurs without update [72,73], which may lead to suboptimal results on new datasets. In online TL the model continuously updates as new data become available– in the target domain, data generation and real-time processing occur dynamically [74] –allowing it to adapt to changing data distributions [76]. Although more challenging, online TL is preferred as biomedical signals often vary in quantity and distribution over time. Combining continuous updates from online learning with knowledge derived from a source domain, online TL can enhance the effectiveness of ongoing tasks in the target domain. It is particularly useful for real-time applications, e.g., driver drowsiness detection [75, 76], continuous epilepsy monitoring [77], BCI [78], and wearable devices [79].

TL approaches can also be classified into four groups: (1) instance-based; (2) feature-based; (3) parameter- or model-based; and (4) relational or adversarial-based [36,54]. Instance-based TL operates on the premise that specific parts of data from the source domain can be leveraged to enhance the effectiveness of target-domain classifiers by allocating and tuning the weights of data from the source domain using methods like reweighting [80] and importance sampling [81]. Feature-based TL creates a feature representation that minimizes discrepancies between the source and target domains while preserving maximum information from both. Methods used can be categorized into two subtypes: (1) asymmetric, which transforms source features to align with target features; and (2) symmetric, which discovers a common latent feature space and transforms source and target features into a new, unified representation [36]. Model- or parameter-based TL involves reusing knowledge within a model by incorporating pre-trained layers in various combinations, with some layers frozen, some fine-tuned, and others newly added. Relational or adversarial-based methods extract transferable features by exploiting logical relationships or rules from the source domain, or by using techniques inspired by GANs [82].

Deep TL is implemented using either weight initialization or fine-tuning [42]. In weight initialization, pre-trained model weights are transferred to a new dataset and updated; in fine-tuning, certain layers remain frozen while others are updated. In the popular fine-tuning method, the only trainable layer is the final classification or fully connected layer, with all other layers frozen. Fine-tuning is particularly useful when the new task is closely similar to the task the model was initially trained on. Before applying fine-tuning, it is important to design a strategy based on several factors, including the quantity of data available, the pre-training datasets used, and the nature of the task being solved. Each factor is critical in deciding the extent of weight freezing and the layers that should be updated during fine-tuning. When dealing with a new dataset that is small and closely related to the pre-trained task, the common method is to retrain only the output layers' weights while keeping the other layers' weights frozen. This method applies if the task is similar and there is adequate data. In contrast, if the new task differs significantly from the pre-trained one and the dataset is limited, the initial layers are kept frozen, and the later layers' weights are retrained with the new data [42,83]. Sometimes, all weights may need to be updated with the new dataset to adapt to the new task effectively.

For the analysis of biomedical signals, TL methods primarily involve feature-based and model-based approaches, as variability between subjects, sessions, and domains complicates the generalization process for models. Feature-based TL mitigates the discrepancy between the source and target domains by extracting unchanging features using techniques like domain adaptation [66,84] and maximum mean discrepancy (MMD) [85]to adjust feature distributions, thereby facilitating more efficient knowledge transfer across different tasks, sessions, subjects, or domains. In model-based TL, models are either employed as feature extractors or adapted for specific analysis tasks. Convolutional neural network (CNN) models pre-trained on 2D images into ImageNet dataset must be adapted for 1D time-series biomedical signal data like EEG and ECG. To apply these models effectively, the signals must be converted into 2D representations, which may be achieved by transforming the time-series signals into spectrograms or scalograms via techniques like short-time Fourier transform (STFT) or wavelet transform (WT).

3. Review methodology

We conducted a systematic review on the utilization of TL in analyzing biomedical signals for various health-related applications, including the brain, heart, and BCI applications per PRISMA guidelines [86]

3.1. Search strategy and selection criteria

We searched PubMed, Web of Science, Institute of Electrical and Electronics Engineers (IEEE), Xplore Digital Library, and Scopus databases for papers published up to 30 October 2024 using Boolean combinations of various keywords related to TL-based medical applications and the relevant biomedical signals employed: "transfer learning, '' "pre-trained, '' "biomedical signals," "machine learning, '' "deep learning, '' "detection, '' "classification, '' "prediction, '' "diagnosis, '' "medical, '' "healthcare, '' "mental, '' "health'', "electroencephalogram'', "EEG'', "electrocardiogram'', "ECG'', "photoplethysmogram'', "PPG'', "galvanic skin response'', "GSR'', "eye tracking'', "ET'', "electrooculogram'', "EOG'', "blood pressure'', "BP'', "polysomnography'', "PSG'', "ballistocardiogram'', "BCG'', "functional

near-infrared spectroscopy'', "fNIRS'', "electromyography'', "EMG'', "magnetoencephalography'', "MEG'', and "multi-modality''.

The initial search yielded 2390 publications, of which 1064 were duplicates. Subsequent screening of article titles and abstracts removed conference papers, non-journal articles, editorials, perspectives, reviews, and articles unrelated to TL or healthcare. Further exclusion of articles published before 1 January 2014, non-English publications, and articles without full-text access, 450 articles remained for full-text review. A final round of review excluded non-Q1 journal articles and articles without adequate performance results of ML or DL models, leaving 239 articles included for analysis (Fig. 3).

3.2. Comparison with existing review papers

A primary contribution of this study is its broad scope: we examine studies utilizing various biomedical signals, e.g., EEG, ECG, PCG, PPG, etc. We also comprehensively review diverse healthcare applications, e.g., disease detection/prediction, BCI systems, etc. Our comprehensive review of TL medical applications across diverse biomedical signals thus distinguishes itself from recent review articles that focused on single signal sources (Fig. 4). Han et al. [87] reviewed pre-trained models for ECG-related medical applications, highlighting challenges and future research directions. Several authors have published on TL applications in brain-computer interfaces (BCI) [88–90]. Khan et al. [91] reviewed studies published between 2014 and 2024 on CNN-based TL models for Alzheimer's disease detection and classification using neuroimaging data [94]. Classical ML and TL approaches to improve generalizability in emotion recognition using EEG signals across diverse settings were reviewed in [92]. Likewise, Li et al. [59] reviewed studies focused on emotion recognition using EEG signals. The authors in [60,93] discussed key methods in TL and examined their real-world applications in EEG signal processing, highlighting challenges and areas for further research. Ray et al. [93] reviewed TL-based human activity recognition, considering various learning approaches as well as aspects of sensing, including wearable, non-wearable, and ambient sensors. The authors in [94] surveyed papers on application of TL techniques for improving digital health services and healthcare outcomes.

4. Results

Biomedical signals are extensively used for detecting and predicting various health conditions, such as schizophrenia [95], heart disease [96], epilepsy [97], mental health [98], sleep stage [99], emotion recognition [100], and stroke [101]. All 239 reviewed papers are summarized in Tables A.1–A.19 (Appendix A), which provide key information for each article, including the datasets utilized for source and target applications, signal modalities, TL models, and evaluation criteria. Also, in this section, we explore TL models and their key features for 1-D signal applications, as well as the use of TL in clinical applications.

4.1. Number of annual papers published in this field

There has been a consistent growth in research papers in the last decade (Fig. 5), which highlights the potential of TL to address data scarcity and domain mismatches, and to enhance model accuracy and efficiency in health-related applications. TL is, therefore, becoming a key area of study within biomedical signals and healthcare research. The most widely used biomedical signals are EEG, ECG, and PCG; some models used multiple modalities (Fig. 5).

4.2. EEG- based TL applications

EEG, which provides a spatiotemporal map of surface electrical activity of the brain, is common studied in neuroscience research, clinical diagnosis (e.g., Alzheimer's disease [102,103], schizophrenia [104], and mild cognitive impairment [105]), and BCI applications (e.g., speech recognition [106], emotion recognition [107], and rehabilitation [108]). Many EEG-based models assume stable signal data distributions. However, EEG signals are inherently non-stationary and exhibit considerable inter-subject and inter-study variabilities due to variations in brain anatomy, signal noise, individual physiological, and psychological states, which limiting generalizability of classifiers across subjects and sessions. To address this limitation, researchers are increasingly employing TL techniques to enhance model performance across diverse populations and recording contexts, aiming to enhance the practical applicability of EEG models in real-world practices. These

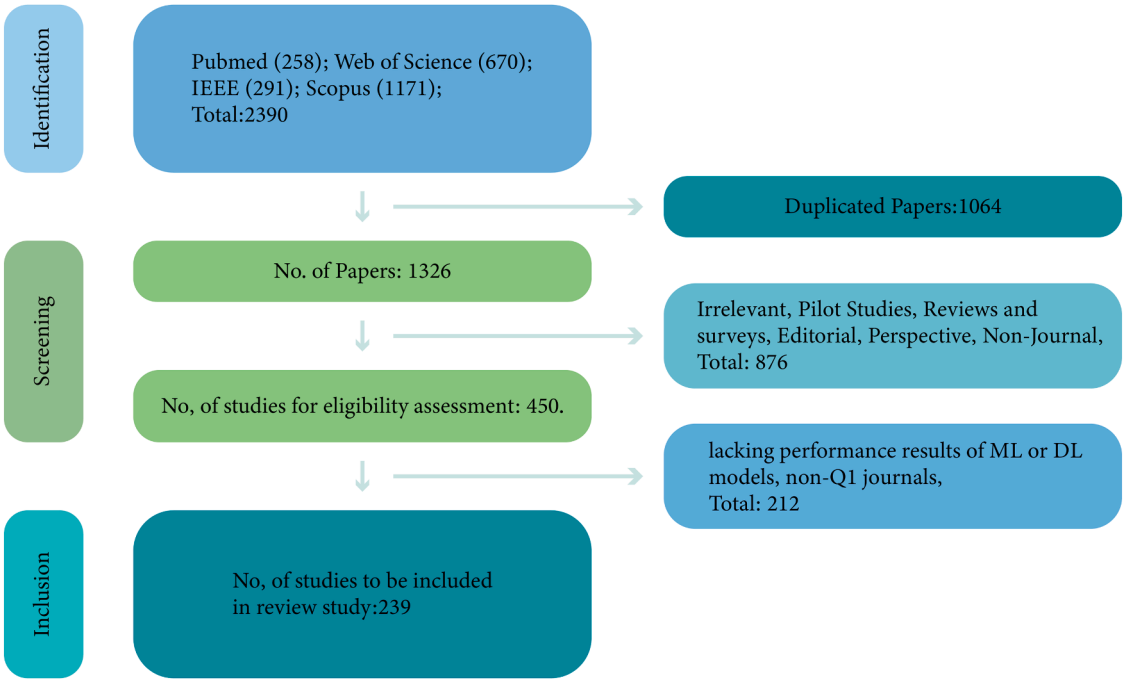

Fig. 3. PRISMA workflow for literature search.

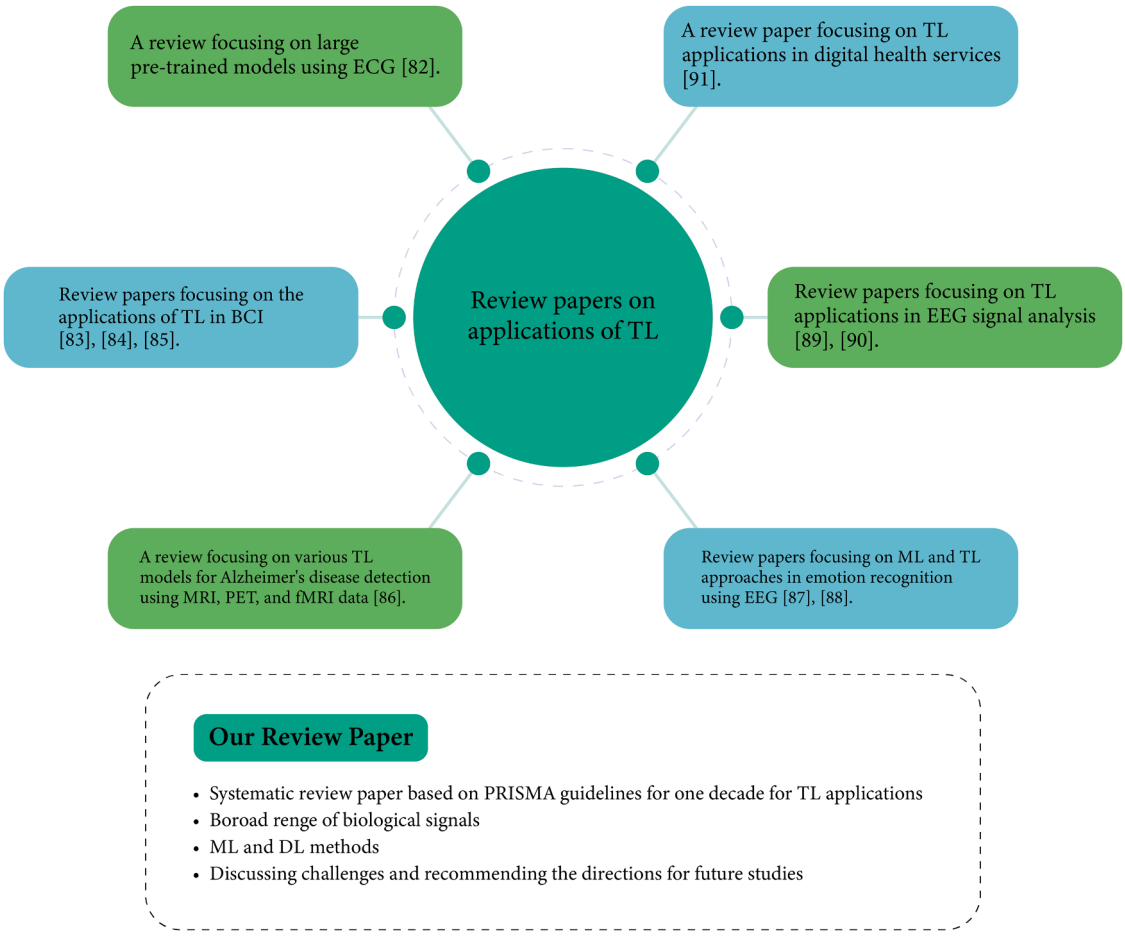

Fig. 4. Comparison of our review paper with other related reviews based on various signals.

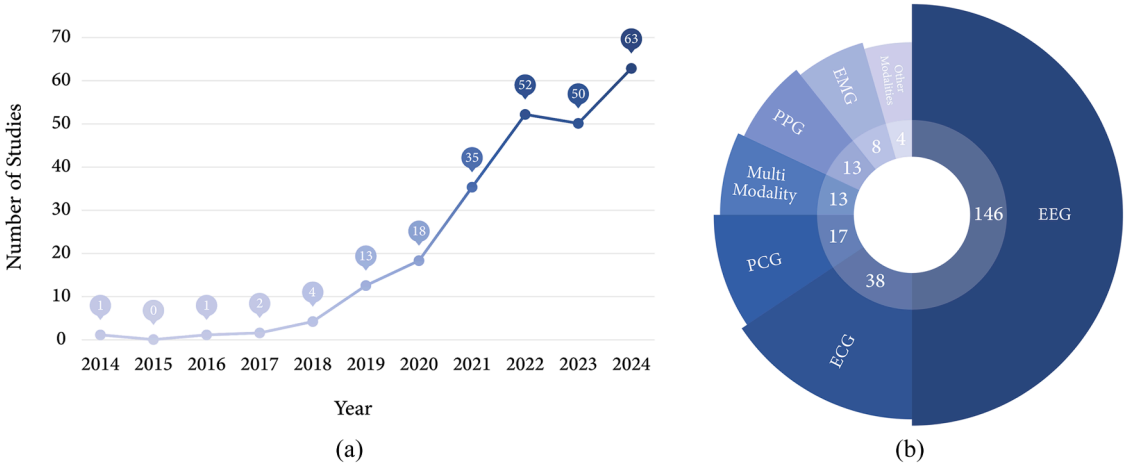

Fig. 5. Secular trend of publications (a) and their distribution by biomedical signal modality (b). Other modalities include galvanic skin response, respiration, skin temperature, and eye tracking (ET).

studies are summarized in Tables A.1–A.9 (Appendix A).

The number of papers utilizing TL methods for EEG research has grown over the past decade, with BCI applications and epilepsy accounting for the largest proportions (Fig. 6), which underscores TL's potential to address key challenges like variability between subjects and tasks, scarce labeled data, and hardware-related constraints. 12 of the 35 studies on epilepsy detection and prediction employed pre-trained models including VGG, ResNet, DenseNet, Inception, and AlexNet (Table A.1), which used 2-D spectrograms or scalograms of EEG signals for epileptic seizure detection. The pre-trained network serves as a fixed feature extractor from the input images or is fine-tuned, partially or fully, to improve epilepsy detection accuracy. While these models, trained on ImageNet, can reduce training time, and achieve good results, they face challenges due to domain mismatches, as features learned from natural images often fail to capture the unique temporal and spatial aspects of EEG signals. Additionally, the inherent variability in EEG

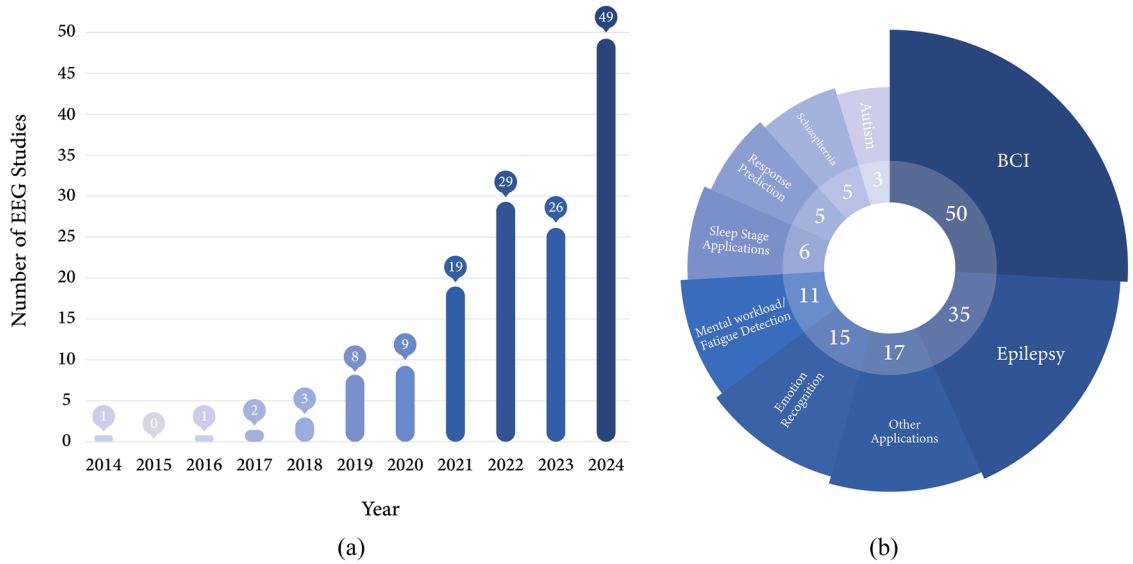

Fig. 6. Secular trend of EEG-related publications (a) and their distribution by medical application (b).

data, such as differences across subjects and recording conditions, further complicates the generalization of these models. These factors often necessitate domain-specific adaptations, or the development of models tailored particularly for EEG data to achieve optimal performance. For example, in [109], the authors used pre-trained models like VGG16, VGG19, ResNet50, InceptionV3, DenseNet121, Xception, NASNet, and ONASNet, and attained best accuracy of 99.67 %. In [110], 99.17 % accuracy was attained for neonatal seizure detection by fine-tuning models like AlexNet, ResNet, GoogLeNet, and DenseNet.

Owing to notable inter-subject variability in EEG signals, most seizure classification models relied on subject-dependent approaches to optimize accuracy. This requires extensive labeling of individual recordings, making it impractical for real-world applications. To address this limitation, researchers have employed TL domain adaptation methods (DA) to transfer knowledge from patient-independent models to patient-specific models, enabling effective generalization across patients. DA for cross-subject EEG classification mitigates data distribution shifts between subjects by aligning features across patient datasets, enhancing generalizability and seizure classification performance while mitigating the requirement for significant labeling of new data. By requiring clinicians to label only a single seizure onset and non-seizure period for a new patient, this TL approach allows for quick fine-tuning of pre-trained models, offering a more efficient and less labor-intensive diagnostic process compared to fully subject-dependent models. While the subject-independent approach is practical for clinical use, only six studies have applied DA techniques to address epileptic-related issues because of prominent inter-subject discrepancies and data scarcity. Cue et al. [111] introduced a subject-specific EEG recognition method that learned domain-invariant features and adapted transfer strategies by comparing feature distributions of source and target samples. Additionally, researchers have focused more on transductive TL than inductive TL, as transductive TL does not require labeled target data. This approach is particularly appropriate for recognizing epileptic EEG signals in scenarios where labeled seizure data in the target domain is limited [112–114].

Five studies used pre-trained models from ImageNet like VGG, ResNet, Inception, DenseNet, EfficientNet, and AlexNet for diagnosis of schizophrenia (Table A.2); three studies used pre-trained models like SqueezeNet, AlexNet, ResNet18, GoogLeNet, MobileNetV2, ShuffleNet, EfficientNet, and DenseNet for diagnosing ASD with satisfactory results (Table A.3). EEG data often exhibit unique characteristics, such as variability in patient demographics, recording conditions, and signal noise, which can introduce biases or hinder performance. These variations between source and target domains may reduce the model's proficiency to generalize effectively. Therefore, improving the similarity between the source and target datasets through techniques like DA or domain-specific fine-tuning provides an effective approach to resolve these issues and improve model performance in medical applications. For instance, in [115], models such as VGG-16, ResNet50V2, InceptionV3, DenseNet121, and EfficientNetB0 were utilized, attaining up to 99.9 % accuracy for schizophrenia diagnosis. In [116], hybrid deep lightweight features extracted from MobileNetV2, ShuffleNet, and SqueezeNet were used for ASD detection.

Eleven studies (8 % of the EEG studies) focused on detecting issues related to mental workload or fatigue (Table A.4). More than half of these studies utilized DA methods to enable cross-subject and cross-task modeling, given the considerable differences in EEG signals among subjects. Wang et al. [117] proposed Sleep EEGNet model (pre-trained on the Sleep-EDF dataset using DA), which facilitated effective cross-domain knowledge transfer, allowing the model to detect fatigue in both virtual and actual situations with 91.5 % accuracy and robust generalizability despite limited data. In [118], the authors introduced an enhanced EasyTL-based technique for cross-subject EEG-based fatigue detection, which yielded satisfactory results, but the efficiency of EEG-based mental workload evaluation across different tasks was limited due to the varying EEG response patterns across varied tasks, which significantly impaired the model's real-world generalizability. In [84,119], the authors applied TL methods to assess cross-task mental workload across various tasks, offering a promising solution to these challenges.

Fifteen studies focused on EEG-based emotion recognition (Table A.5.), a field that faces significant challenges due to differences between individuals in EEG signals to the same stimuli. Much research sought to address performance limitations caused by differences in EEG signals across subjects, sessions, devices, and datasets. Wang et al. [120] introduced a deep multi-source adaptation transfer network (DMATN) to resolve subject-specific data bias and the lack of sufficient training samples for new subjects for online EEG emotion recognition. Ma et al. [85] introduced a cross-subject source domain selection approach to address the accuracy issues in cross-subject EEG-based emotion recognition, which results from negative transfer due to inappropriate data in the source domain. While there has been growing interest in cross-device challenges using TL, few studies have explored cross-device tasks, highlighting an important area for future investigation. Liu et al. [121] applied TL to improve emotion recognition on few-channel EEG data by employing knowledge derived from full-channel EEG data.

A few studies focused on EEG-based prediction of major depressive disorder (MDD) (Table A.6) and sleep stage classification (Table A.7). The five studies related to MDD diagnosis used various pre-trained models and ensembles of these models [122,123]. In [124], the authors attained 96.55 % accuracy using fine-tuned pre-trained models like VGG16, Xception, DenseNet121, MobileNetV2, and InceptionResNetV2 for the target task, and an ensemble method based on majority voting. To address issues of data variability and data inefficiency, the six studies on sleep stage classification enhanced model performance using TL techniques that facilitated knowledge transfer across various scenarios, including cross-channel, cross-subject, and cross-dataset applications. In [125], an unsupervised DA network was successfully applied for sleep staging across channels, subjects, and datasets, enhancing its suitability for practical, everyday applications.

Fifty studies (34 % of the total studies) have focused on BCI applications (Table A.8). To address challenges across various cross-scenarios, DA, which utilizes data from source domains or subjects to aid in calibrating of a new target domain, was widely applied. Deep DA can tackle domain shifts caused by individual differences, where neural responses vary across subjects for the same stimulus. Most approaches used a domain discriminator and adversarial training inspired by GANs to identify domain-invariant features [126,127]. Researchers have increasingly focused on supervised DA, where all the target data labels are used during training, as well as unsupervised or semisupervised DA, where only some or none of the target data labels are available [128, 129]. Additionally, domain generalization involves not using target data during the training process, which helps improve generalizability across subjects and sessions [130]. For example, Roy et al. [131] proposed a TL-based, multi-scale feature-fused CNN handle inter-individual differences in EEG signals, which excelled at identifying distinctive features across different non-overlapping canonical frequency bands of EEG signals, utilizing various convolutional scales for multi-class motor imagery classification. In [132], the authors presented a multi-source TL framework employed optimal transport feature selection for EEG-based motor imagery classification, which attained good 85.93 % accuracy. Zhang et al. [133] developed a semi-supervised multi-source TL model for cross-subject motor imagery-EEG classification, which learned domain-invariant features and incorporated dynamic weighting to integrate features from multiple sources for accurate predictions.

Research is limited in cross-dataset TL, a promising approach that leverages existing motor imagery or BCI datasets to improve model effectiveness on new datasets and reduce the need for intensive dataset-specific data collection. In [134], a cross-dataset TL method based on multi-task learning and pre-training for motor imagery classification was proposed. The authors pre-trained DL models using a source dataset to optimize the models' effectiveness on a target dataset and then fine-tuned the pre-trained models using the target dataset. Miao et al. [135]. introduced a multi-source deep DA ensemble framework to enhance cross-dataset motor imagery EEG decoding. Developing effective cross-dataset TL methods can potentially create more robust and adaptable models for real-world BCI applications that can perform across diverse datasets.

Seventeen studies focused on miscellaneous applications (Table A.9), including EEG-based authentication [136], insomnia detection [137], Parkinson's disease detection [138] diagnosis of hearing deficiency [139], and investigating and identifying abnormal patterns in Alzheimer's disease [140]. Most of these studies utilize popular pre-trained models including VGG, ResNet, and AlexNet, and reported satisfactory results.

4.3. ECG- based TL applications

ECG continuously records the electrical activity of the heart using electrodes positioned on the patient's body. It supplies information on heart structure and rhythm and is commonly used to diagnose diverse heart conditions like arrhythmia, myocardial infarction, and coronary heart disease. The application of DL in ECG diagnosis is limited by the lack of large well-labeled datasets, making TL a useful alternative. 38 studies were reviewed (Tables A.10–A.13). The number of publications has steadily increased, with the majority of the studies dedicated to arrhythmia or atrial fibrillation (AF) detection (Fig. 7). Most of the studies have applied models pre-trained on ImageNet, followed by fine-tuning on ECG datasets. While these models yield promising results, they face challenges in generalizability and practicality due to domain mismatches, limited labeled data, and the need for domain-specific solutions. A few studies have applied DA methods to mitigate domain shifts in ECG classification. Individual differences, such as age, sex, physiology, and recording devices, complicate feature learning. In DA, training and test samples are utilized as source and target domains, respectively, aiming to learn shared features despite differing feature spaces. However, traditional DA frameworks often neglect data structure and semantic conformity, which can reduce effectiveness or cause negative transfer. Unsupervised DA is a promising approach that can eliminate the need for labeled target data, thereby reducing expert labeling and maintenance costs.

Ten studies focused on arrhythmia detection (Table A.10). Detection of arrhythmia (abnormal heart rhythm) caused by abnormal cardiac electrical impulses and conduction is clinically important, but limited data availability and significant data imbalance impose hurdles to model development. Collecting sufficient data for rare arrhythmia types is difficult, which makes identifying minority cardiac arrhythmia particularly challenging. 7 out of the 10 studies utilized models pre-trained on ImageNet, including ResNet50, AlexNet, SqueezeNet, DenseNet, GoogleNet, VGG-16, Inception-v3, and MobileNet, which were subsequently fine-tuned using ECG datasets [141,142]. In [143], the authors proposed a multi-level unsupervised DA method for inter-subject arrhythmia diagnosis.

AF is a specific and the most prevalent arrhythmia, which is associated with morbidity and mortality that can be prevented with early detection and appropriate treatment. The studies are summarized in Table A.11. In [144], an unsupervised DA technique was employed, which used small and significant volumes of labeled and unlabeled data, respectively, for training. Ng et al. [145] introduced a personalized AF detector based on a Siamese network, which used few-shot learning to address the issue of imbalanced datasets during the fine-tuning process. In [146], the authors employed TL for the automated detection of AF among patients being evaluated for suspected obstructive sleep apnea. They used a pre-existing ECG model on single-lead ECG traces recorded during in-laboratory sleep studies, without the need for further training of the algorithm. Xu et al. [147] developed a dynamic DA approach to reduce the influence of distribution discrepancies through adaptive learning of ECG features from both source and target domains for cross-database AF detection.

Eight studies focused on other cardiovascular diseases, including myocardial infarction (MI), genetic heart diseases, cardiac anomalies, etc. (Table A.12) Five of these studies used pre-trained models from ImageNet for knowledge transfer, followed by fine-tuning on ECG datasets [148]. A few studies employed models pre-trained on large ECG datasets, which were then fine-tuned using the target datasets. Pre-training on a large, similar ECG dataset helps the model learn detailed and robust signal features. The optimized parameters from pre-training are transferred to the target dataset, enabling the model to build on the rich ECG representations it learned, establishing a reliable foundation for reliable tasks [149,150].

Ten studies focused ECG-based diagnosis of non-cardiovascular diseases (Table A.13), including stress prediction [151], diabetes mellitus (DM) detection [152], emotion recognition [153], noise detection and classification [154], and COVID-19 detection [155]. Eight of these studies used pre-trained models like Xception, GoogLeNet, DarkNet-53, ResNet, InceptionResNetV2, DenseNet, InceptionV3, VGG, and AlexNet, which were subsequently fine-tuned on ECG datasets to adapt them for specific applications. Srivastava et al. [156] introduced an ensemble of

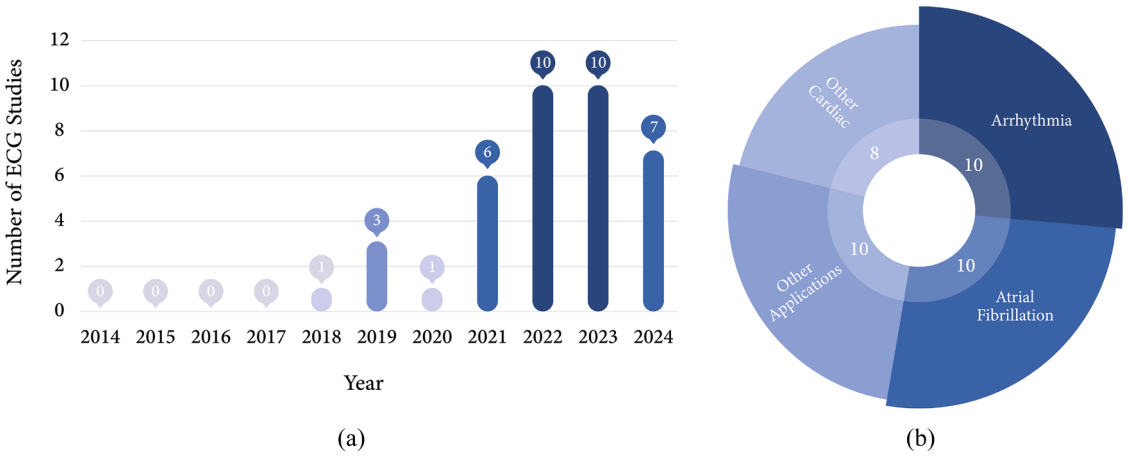

Fig. 7. Secular trend of ECG-based publications (a) and their distribution by medical application (b).

pre-trained DNNs, including ResNet and DenseNet, for ECG-based biometric recognition and reported promising results. In [157], the scalogram of the ECG was utilized as input to pre-trained models, including AlexNet, GoogLeNet, and ResNet, for biometric classification.

4.4. PPG-based transfer learning

High blood pressure (BP) is a prevalent public health concern that can cause brain and kidney end-organ damage [158], making early detection and treatment important [159]. PPG, a noninvasive method that measures changes in blood volume within tissues using photoelectric technology has been used for detecting and monitoring BP as well as arrhythmia [160], including on wearable PPG sensors. Thirteen studies focused on PPG-based research (Tables A.14 to A.15). The first study on PPG-based TL research was published only in 2020, with most focusing on BP estimation (Fig. 8) [161]. In [162], the authors employed large source PPG datasets that were substantially different from the target datasets to enhance the robustness of models for BP estimation across diverse datasets. Qin et al. [163] employed a deep generative model combined with domain adversarial training to address the challenges associated with individual variability by extracting features that remain consistent across different individuals. In [164], the authors proposed a self-supervised TL model using transformer model to extract strong representations of transformed PPG signals in the pre-training phase. Additionally, the TL approach included BP pattern adaptation to identify distinctive features for precise BP value estimation.

Among studies that focused on PPG-based TL for other applications (Table A.15), Song et al. [165] developed a remote Photopletysmography (rPPG) method for estimating heart rate from facial videos. Their approach uses CNNs to map spatio-temporal physiological features to their corresponding ground truth, with the ResNet-18 network enhancing HR prediction accuracy. Recent research has highlighted the effectiveness of PPG analysis in conducting extensive screenings for detection of diabetes. Zanelli et al. [166] utilized a Light CNN model to detect type 2 diabetes based on a single raw pulse derived from PPG signals. They also implemented TL, initially training the model on a large dataset before fine-tuning it on a smaller dataset of PPG signals specific to type 2 diabetes. They also investigated the relationship between hypertension and diabetes. Osathitporn et al. [161] introduced RRWaveNet, which employed multi-scale convolution and residual CNN to analyze PPG data in a subject-independent manner. To tackle challenges of limited data and variations in data distribution across datasets and devices, they applied a TL approach.

4.5. PCG-based transfer learning

PCG records heart sounds, which can be used to diagnose various heart conditions [167]. Thirteen studies focused on PCG-based diagnosis of various heart conditions (Table A.16). The first article was published only in 2020 (Fig. 9). In many real-world situations, PPG has variable distributions that are frequently imbalanced, leading to unequal representation of classes. By decreasing the majority class's influence and

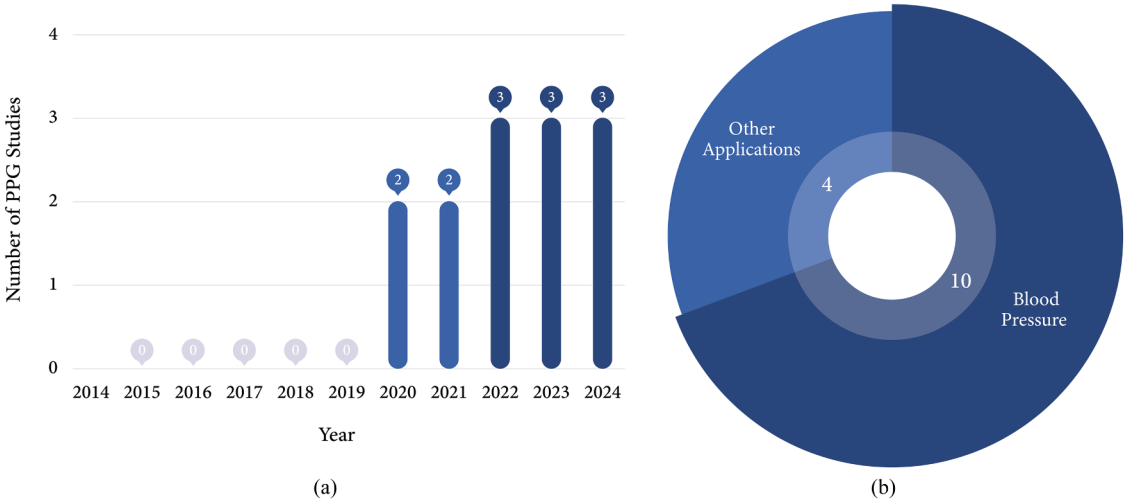

Fig. 8. Secular of PPG-based publications (a) and their distribution by application (b).

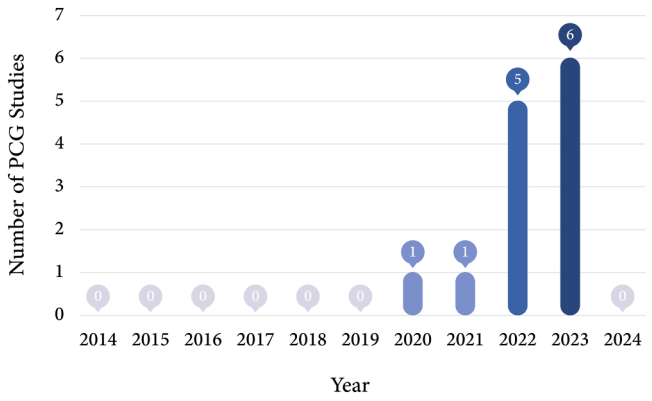

Fig. 9. Distribution of published papers over time using TL technique and PCG signals.

increasing the emphasis on the minority class during the training process, ensemble models can effectively tackle this issue. Singh et al. [168] introduced an ensemble-based TL approach using models likes AlexNet, SqueezeNet, and VGG19, along with spectrogram images, to resolve the data imbalance in PCG classification. In [169], the authors proposed an automated approach for diagnosing heart valve diseases using a CNN and time-frequency-domain images of PCG signals and benchmarked their results against time–frequency domain TL models like ResNet-50 and VGGNet-16 and demonstrated that the developed deep CNN outperformed both, achieving higher overall accuracy. Zheng et al. [170] applied pre-trained models like VGG16, Xception, ResNet50, and InceptionResNet-V2 to extract a range of domain-specific deep features from PCG spectrograms. These selected features were subsequently combined and input into a CatBoost model to classify cardiac diastolic dysfunction.

4.6. Multi-modality-based transfer learning

Thirteen studies focused on cross-domain TL [171], a variant of TL, that merges knowledge from different domains or modalities, such as bridging multiple biomedical signals from different medical imaging inputs or combining medical imaging insights with clinical data (Table A.17), e.g. refining fMRI scan models using EEG data insights [172]. This approach can lead to more comprehensive and accurate predictions. Owing to the heterogeneous nature of medical data, which includes clinical, imaging, and genetic information, cross-domain TL plays an integral role in advancing our knowledge of diseases [173], and mimics the way neuroscientists manually integrate diverse information from multimodality neuroimaging, including various brain structural and functional modalities [174–176]. The first article was only published in 2019; the studies involved multiple signal types (e.g., EEG, ECG, PPG) and diverse applications, including heart disease detection, sleep stage classification, and emotion recognition.

Biomedical signals such as EEG and ECG exhibit significant variability across populations and settings, posing challenges for cross-domain transfer and model reliability. Karthikeyan et al. [177] proposed a cross-modal TL framework that integrates clinical records, medical imagery, and genetic information, leveraging attention mechanisms to predict cardiovascular disease. Their approach achieved outstanding performance, with 93.5 % accuracy, 92.0 % precision, and 94.5 % recall. In [178], researchers classified sleep stages using features extracted from EEG and EOG signals. In [179], the authors developed a multi-modal EEG-EOG classification system that included a VGG network trained from scratch, as well as decision-making systems for sleep staging. Giovannetti et al. [180] proposed a framework for detecting early signs of Alzheimer's disease by combining longitudinal Magnetoencephalography (MEG) data with brain MRI using a pre-trained AlexNet to generate functional connectivity maps. In [181], the authors developed two DL-based strategies to identify cognitive workload during surgical tasks using multimodal signals, including EEG, fNIRS, and pupil eye data PE, employing a pre-trained AlexNet to detect atypical cognitive workload levels. In a second strategy, a 1D convolutional neural network was utilized, with a 1D vector array formed by concatenating EEG, fNIRS, and PE data serving as the input for training. Fig. 10 illustrates the distribution of published papers employing the TL technique and multimodal signals over the past decade.

4.7. EMG-based transfer learning

EMG, which captures electrical signals generated by muscles during contraction, reflects the anatomical and physiological characteristics of the muscle motor unit [182]. EMG signals are used in medical diagnosis of muscle disease. EMG signals can be acquired by surface EMG (sEMG) or intramuscular EMG (iEMG) electrode [183]. Non-invasive sEMG), collected from the skin surface, is inexpensive and easy to obtain, iEMG is invasive. Eight studies focused on EMG (Table A.18), with sEMG being used mostly used for gesture classification (Fig. 11).

Three studies used models pre-trained on ImageNet [184–186]. However, the significant disparity between natural images of the source domain with spectral images of EMG signals in the target domain raises questions about the extent to which knowledge is transferable between such distinct domains. Phoo et al. [187] explained that augmenting the data does not necessarily lead to improved generalizability, as no additional informative knowledge can be extracted from the source domain to benefit the learner in the target domain. Due to domain shifts in EMG signals, the application of TL methods, including various domain adaptation models, is crucial for addressing shifts between individuals, sessions, and modalities. This approach significantly enhances the generalizability of the model. Furthermore, research should focus on real-world scenarios, such as online experiments, to confirm these models can be effectively deployed in real-world scenarios like wearable devices. In [188], the authors employed a TL approach with DA to tackle cross-day and cross-subject variability in hand gesture recognition using sEMG. Additionally, in [189], the authors propose a new virtual reality experimental protocol that uses an adaptive domain adversarial neural network for sEMG-based gesture recognition. Adversarial-based methods focus on encouraging the neural network to learn hidden EMG representations free of domain-specific differentiating information, achieved through an adversarial training process.

4.8. Other modalities-based transfer learning

Four studies focused on use of miscellaneous signals, i.e., heart rate, polysomnogram (PSG), ballistocardiography (BCG), and functional

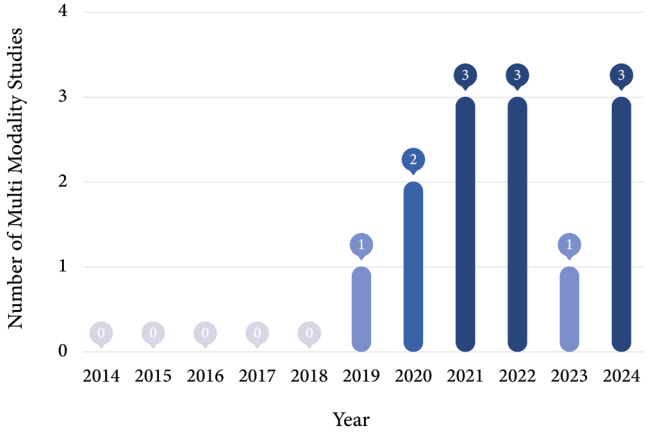

Fig. 10. Distribution of published papers using TL technique and multimodal signals over time.

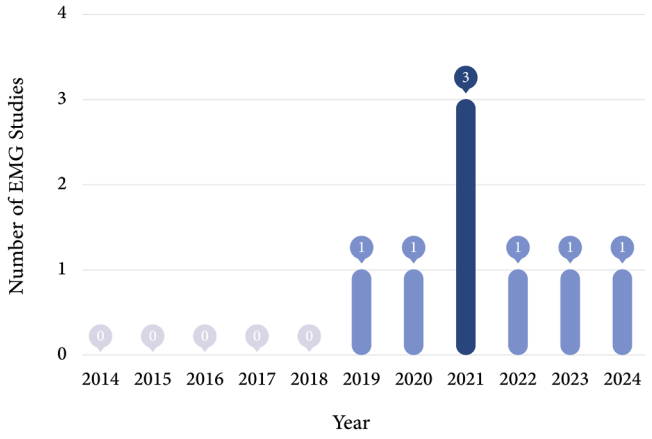

Fig. 11. Distribution of published papers using TL technique and EMG signals over time.

near-infrared spectroscopy (fNIRS) (Table A.19, Fig. 12). PSG, which is collected during an entire night of sleep, is the gold standard for diagnosing sleep apnea. Because PSG is often cumbersome, expensive, and time-intensive, various solutions have been developed to streamline sleep scoring by automating the process. These include using a single-lead ECG, monitoring pulse rate (PR) with PPG and employing BCG [190]. In [191], authors proposed an automated hypertension detection system that employs time–frequency (T-F) spectral images and CNNs to precisely identify hypertension using BCG signals. The T-F spectral images were analyzed using AlexNet, ResNet-50, and the newly developed Hyp-Net CNN, all validated through a 10-fold cross-validation method. fNIRS, a non-invasive functional neuroimaging technique, serves as a safe and portable method for assessing the cortical control of gait. Ma et al. [192] proposed an automatic method for classifying walking tasks that require diverse levels of cognitive resources using fNIRS data. They fine-tuned pre-trained models, including ResNet, VGG, MobileNet, EfficientNet, and TinyNet, on the fNIRS dataset, achieving an accuracy of up to 81 %. Fig. 12 presents the distribution of published papers utilizing the TL technique alongside other modalities over the past decade.

4.9. TL models and key features used for 1-D signal applications

To conduct this systematic review, we examined 239 studies from the past decade on the applications of TL in biomedical signals. The majority of the studies utilizing TL were related to EEG and ECG modalities, with 146 and 38 studies, respectively. Table 1 summarizes the most commonly used architectures in TL applications and key features for EEG and ECG signals. Among EEG applications, brain-computer interfaces (BCIs) and epilepsy detection have garnered the most attention.

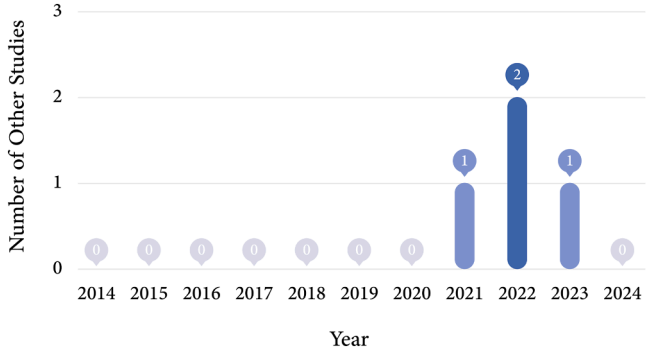

Fig. 12. Distribution of published papers using TL technique and other modalities over the past decade.

Table 1
Summary of commonly used TL models and key features for EEG and ECG signal applications.

Modality	Application	Model	Key Features	Variants
EEG	BCI	DA	**DA**: Effectively handles domain shifts and individual differences, improving feature extraction and generalization across diverse domains and enhances the ability to extract domain-invariant features using domain discriminators and adversarial training.	Supervised, Semi-supervised, Unsupervised, Adversarial DA
	Epilepsy	ResNet	**ResNet**: Deep architecture with residual blocks with skip connections between layers, addresses vanishing gradient problem, and is scalable to deeper versions [193].	ResNet-50, ResNet-101
		VGG	**VGG**: Simple 3 × 3 convolutional filter-based architecture, deep yet computationally efficient, excels at learning hierarchical features [194].	VGG-16, VGG-19
		Inception	**Inception**: Uses parallel convolutions with various kernel sizes (1 × 1, 3 × 3, 5 × 5) within inception modules, replacing large convolution layers for spatial filtering, reducing dimensionality through 1 × 1 convolutions, and enhancing the network's nonlinearity and depth [195].	Inception v1, Inception v3
		AlexNet	**AlexNet**: Shallow architecture uses large kernels and overlapping max-pooling to reduce dimensionality while retaining critical information [196].	AlexNet
ECG	Cardiac (arrhythmia, atrial fibrillation, other cardiovascular)	ResNet		ResNet-50, ResNet 152, ResNet-34v2, ResNet-101
		VGG		VGG-19, VGG-16
		AlexNet		AlexNet

In the domain of epilepsy detection, specific pre-trained models, including ResNet, VGG, AlexNet, and Inception, have gained considerable popularity due to their good performance in addressing common issues in convolutional networks, such as vanishing gradients, degradation, long training times, and the excessive number of trainable parameters and lack of data. These models are pre-trained on ImageNet, which differs significantly from biomedical data, so the training strategy involves fine-tuning by freezing the initial layers and training the weights in the final layers, with features often extracted from the top layers to enhance the model's ability to generalize to the new data. If the model's performance requires refinement, it should be fine-tuned by progressively unfreezing the convolutional layers from the top to the bottom, using a low learning rate. For BCI applications, Domain Adaptation (DA) methods are extensively employed to tackle challenges across various scenarios by utilizing data from source domains or subjects to calibrate new target domains. This approach facilitates effective management of domain shifts and enhances feature extraction across diverse domains. The focus is on supervised DA, which involves using target labels during training, as well as unsupervised or semi-supervised DA, where partial or no target labels are utilized.

Another widely studied modality using TL is ECG. Among the 38 studies reviewed, 28 focused explicitly on cardiovascular diseases, representing the most extensively researched area within this field. Most ECG applications utilize pre-trained models and have achieved promising results. In some studies, an ensemble of pre-trained models yielded good results, while others relied solely on pre-trained models for their applications.

4.10. TL in clinical applications

In clinical settings, applying TL-based solutions holds great potential to enhance diagnostic accuracy, treatment planning, and patient monitoring. However, translating these solutions into practice requires overcoming several challenges. One key concern is data bias, which can arise from humans' interpretation and selection of data. In healthcare, this could lead to models that perform well in specific populations but fail to generalize across diverse groups of patients. Therefore, ensuring that datasets used for training models are comprehensive and diverse is critical. In the medical field, the limited availability of data presents a challenge; therefore, accessing data from multiple centers is essential to ensure the model's representativeness and robustness while also maintaining patient privacy and complying with regulatory requirements. Another significant consideration is the trust and confidence of clinicians in AI models. To foster this trust, it is essential to integrate explainable AI (XAI) techniques, which provide transparency into how models arrive at decisions. This helps clinicians understand the rationale behind the model's predictions, making them more likely to adopt and rely on these tools in clinical practice. Additionally, uncertainty quantification is crucial for assessing the reliability of model predictions. By providing confidence in the model's output, clinicians can determine when to act on a prediction and when further verification is needed, ensuring fairness, inclusivity, and more informed decision-making during model deployment [197]. It is also essential to consider the practical aspects of implementing these models, including integration into existing clinical workflows and addressing concerns about reliability, accountability, and interpretability. To navigate these challenges effectively, the development of robust regulatory frameworks is essential. For example, a review of Ghana's Public Health Act underscores the importance of algorithmic accountability, safety audits, and updated regulations to address legislative gaps, ensuring that AI-driven solutions are safe and ethical for healthcare settings [198]. Ensuring the successful implementation of TL-based solutions in real-world healthcare requires a multidisciplinary approach involving regulatory bodies, healthcare providers, and technology developers to align on ethical standards, transparency, and patient safety.

5. Public datasets

Access to large, high-quality public datasets is crucial for training models in biomedical signal analysis, especially DL models, which require substantial data to improve their analytical capabilities and overall performance. Large datasets improve model generalization by exposing them to a wide array of patterns and variations, improving their performance on new data. As collecting and annotating biomedical signals is resource intensive, data scarcity is a universal challenge in this field. By reusing existing data for pre-training, models can learn foundational features that enhance their performance when fine-tuned on smaller, specific datasets. Public datasets also support robust benchmarking, providing a standard for evaluating methods and fostering collaboration within the research community. They help to prevent overfitting by ensuring models have enough data to learn generalizable representations. Additionally, open-access datasets encourage interdisciplinary collaboration between biomedical and computational fields, which is vital for advancing research in biomedical signal analysis. Increasing the availability of these datasets is essential for driving innovation and progress in this area. We have categorized publicly available datasets used in studies of biomedical signal analysis based on their applications in healthcare in Table D.22 (Appendix D). We highlight key features of each dataset and provide links for accessing them.

6. Challenges and future directions

The advantages of TL over traditional approaches for biomedical signals are evident: it requires only a small sample of signals from the target domain to develop effective models, thanks to the knowledge transferred from the source domain. Additionally, TL significantly improves the generalization ability of models by reducing the distribution gap caused by substantial data variability and noise interference. Despite these benefits, several methodological and data-related challenges still need to be addressed within the current TL framework. The challenges in applying TL to biomedical signals include dataset limitations, issues with TL models, negative TL, explainability, and hardware constraints. Overcoming these challenges is fundamental for enabling practical uses such as personalized healthcare, wearable devices for continuous monitoring, and BCI systems. To achieve this, we recommend improving access to large, balanced, and multimodal datasets, advancing online TL for adaptive real-time scenarios, enhancing explainable AI for clinical trust, refining DA techniques to address cross-domain variability, exploring few-shot learning and transformer models for efficient learning, addressing hardware limitations for scalable deployment, and leveraging federated learning for secure and collaborative model updates. These steps are vital for translating TL methodologies into impactful, practical solutions in healthcare.

6.1. Challenges in datasets

As mentioned, TL methods typically require extensive data for pre-training; however, publicly available datasets, which incorporate various medical modalities, often involve a limited number of subjects. The reluctance of patients with different health issues to participate in modality recording, ethical concerns, and a lack of confidence among physicians in CADs complicate the provision of comprehensive medical datasets to large populations. It is also worth noting that some researchers have relied on clinical datasets collected through different methods. As a result, these datasets may adhere to varying recording standards, making comparative studies in this field challenging. Furthermore, the class imbalance is a prevalent issue in medical applications; for example, the positive class (representing an event or condition) may have fewer instances, while the negative class (indicating the absence of an event or condition) has a larger number of instances, which can adversely impact detection accuracy. Furthermore, the need for more multimodal datasets with a significant number of subjects

poses another challenge. Access to such datasets would enable researchers to develop sophisticated artificial intelligence approaches that assist medical professionals in the timely diagnosis of diseases.

6.1.1. Limited availability of large datasets

The necessity for extensive datasets for pre-training TL models arises from the requirement for these models to learn detailed features and knowledge, which enhances their capacity to generalize to new, unseen samples. In medical applications, the limited availability of large datasets exacerbates the challenge of limited data for pre-training, and relying on extensive datasets such as ImageNet is problematic. These variations between source and target datasets can result in suboptimal performance, increasing the likelihood of bias and domain mismatch [199]. Many studies utilized pre-trained models on ImageNet, fine-tuning them with biomedical signal datasets. A logical approach to mitigate these issues would be to enhance the alignment between source and target datasets. Therefore, careful selection of pre-training datasets is essential to prevent domain mismatches, as failing to meet necessary standards can create noise and lead to negative transfer, ultimately reducing the overall performance of the TL method [200,201]. However, assembling a large dataset of medical data is challenging due to the complex process of data labeling by expert doctors, difficulties in accessing certain medical data for disease diagnosis, and medical ethics considerations. Therefore, overcoming these challenges could improve the performance of TL models for health-related applications. As mentioned, the primary objective of this study is to examine articles on TL applications for biomedical signals across various applications. Therefore, future efforts could focus on providing public medical datasets with a large number of subjects for pre-training models in a subject-independent and domain-specific way, enhancing their accuracy and generalizability. This marks an important step in healthcare by using AI to make the clinician's workflow faster and more efficient.

6.1.2. Class imbalance

In medical settings, many datasets experience class imbalance, making it difficult to identify rare conditions. This challenge occurs because many classification algorithms presume an even distribution of classes, which leads to a bias towards the class with the most samples. Collecting a balanced dataset for training is particularly challenging in medical and clinical settings due to the low prevalence of certain diseases. This imbalance means that the number of samples among various classes is uneven, which can adversely impact model performance and reduce the generalization of these models. Models that excel with balanced datasets might encounter difficulties when dealing with imbalanced datasets [202]. While class imbalance is a common issue in clinical practice, most modern AI models require balanced datasets to achieve optimal performance. A key challenge with class imbalance is that models trained on unbalanced datasets may be overfit to the less represented classes and show inferior performance during testing [202]. To address this issue, various strategies are employed at both the data and algorithmic levels. Data-level techniques include methods like synthetic data generation [203,204], rebalancing [205], and resampling [206] which aim to enhance the representation of smaller classes in the dataset. On the other hand, algorithm-level methods like cost-sensitive learning [207], class-balanced loss [208], weighted loss functions [209] aim to balance the learning process by assigning higher weights to loss functions for classes with fewer samples. This approach imposes unequal penalties for classification errors, ensuring that under-represented classes receive greater attention during training, ultimately improving model performance on imbalanced datasets. Additionally, DL solutions like generative adversarial network-based methods and autoencoders are often recommended. However, models trained on synthetic data frequently need help generalizing to real-world data and may exhibit lower performance. Future efforts should be directed toward developing advanced data augmentation and synthesis techniques and dynamic and adaptive resampling methods architectures to improve the performance of imbalanced datasets.

6.1.3. Multimodality

Specialist doctors use multimodal data, such as the integration of structural and functional neuroimaging modalities or the combination of biomedical signals/images with patients' clinical records, to accurately diagnose diseases that have complex symptoms. TL in multimodal settings faces several challenges, including variations in feature spaces and data representations across different modalities, the absence of a shared feature space for effective integration, and the potential for negative transfer when knowledge from unrelated source domains adversely affects the target domain [210]. Developing algorithms to align and map representations between modalities while preserving key features is also difficult. To tackle these issues, it is necessary to develop novel fusion techniques and domain adaptation strategies. Having access to larger multimodal datasets can function as a useful resource for advancing diagnostic applications in medicine. Conventional diagnostic approaches typically depend on a single modality, such as cognitive assessments, medical imaging, or biomedical signals. Each method offers only a partial perspective on the disease, whereas combining multiple methods can yield a more thorough and accurate evaluation. By integrating diverse data sources, multimodal approaches leverage the complementary strengths of each modality, improve spatial and temporal resolution in understanding disease conditions, and enhance data quality across modalities [211]. Therefore, developing models that can generalize knowledge across different medical data types, e.g., transferring learned features from ECG to PPG or from EEG to MEG, assists doctors and researchers in recognizing patterns, correlations, and anomalies crucial for diagnosing diseases or distinguishing them from other conditions.

6.2. Negative TL

Negative transfer happens when knowledge acquired from a source domain negatively impacts performance in a target domain, often because of the variation in data or an inefficient transfer method. This issue arises when a mismatch occurs between the source and target data or when the transfer technique is unable to identify the useful parts of the data [212]. To prevent this, it is crucial to assess how transferable the data is between different domains or tasks and to evaluate their similarities before developing models. Negative transfer can also occur if all the source data is utilized to predict labels in the target domain, even though only a portion of the source data is relevant. In such cases, it is important to analyze the transferability of the data before calculating any distances to avoid the negative impact of irrelevant information. Negative knowledge transfer is a frequent challenge in TL methods applied to biomedical signals due to the substantial domain gap, high data complexity and variability, and significant subjectivity in modeling. In the medical field, domain mismatch arises when significant discrepancies exist between the source and target domains regarding subject, session, modality, different labeling functions and channel count. For instance, while pre-training datasets like ImageNet predominantly feature natural 2D RGB images, specialized fields like medicine may include diverse data types including 2D grayscale images, 1D time series, 3D structural data, or video data with accompanying time series information [213]. Variations in color, features, and pixel-level details can heighten the likelihood of discrepancies between the source and target domains, resulting in particular difficulties in designing robust models for medical applications [71]. As a result, researchers are driven to develop and refine TL-based approaches, such as DA [69,214]and domain generalization [130], which are designed to effectively manage these variations and enable the application of pre-trained models across diverse scenarios.

Domain adaptation has been recognized as an effective strategy to tackle this issue, focusing on reducing the distribution gap between different but related domains. It has garnered considerable attention

over the past few years. DA strategies can be categorized into two types according to the nature of distribution differences between source and target domains: marginal distribution adaptation [215,216]and conditional distribution adaptation [217,218]. Marginal distribution adaptation has seen significant progress in various cross-subject tasks and multi-task learning scenarios where multiple source domains correspond to a single target domain. In contrast, conditional distribution adaptation addresses knowledge transfers when the conditional distributions differ between the source and target domains. Research in this domain is less developed compared to marginal distribution adaptation due to its higher complexity. Currently, DA primarily concentrates on offline learning for marginal distribution adaptation. Moving forward, exploring online joint distribution adaptation, which handles both marginal and conditional distribution differences, could provide a more comprehensive understanding of how these variations affect classification results. To further expand and improve the potential of TL, future research should prioritize the development of advanced DA methods, including unsupervised and semi-supervised DA [214,219,220], as well as meta-learning techniques. These efforts aim to address distributional differences across subjects, tasks, sessions, and datasets in various cross-domain scenarios enabling systems to learn effectively from limited patient data. Additionally, there is potential to investigate cross-device tasks using TL, which has been scarcely explored. This approach could enable the processing of signals from portable and compact devices, facilitating the widespread use of biomedical signals in everyday applications.

6.3. Ethical considerations

Privacy protection is of utmost importance in the medical field and typically prevents the sharing of patient data. Therefore, obtaining large-scale labeled multicenter datasets in the medical field is often constrained by data-sharing restrictions and ethical considerations. Federated learning (FL) addresses these issues by offering a solution that respects data privacy and ethical standards [221,222]. This method enables the collaborative training of models across healthcare institutions while avoiding the need to exchange sensitive data. Instead, each device trains its model locally and transmits the network weights to a central server, where they are combined to create a comprehensive model [223,224]. This approach enhances the global model's detection capabilities by incorporating additional insights from multiple sites, effectively addressing the problem of limited supervision associated with training AI models on small datasets [223,224]. Despite its conceptual advantages, FL still needs several practical limitations, such as data heterogeneity. This is because of the differing distributions between the data of specific users and the shared model. Future research into FL methods for TL using biomedical signals should focus on addressing domain shifts such as inter-subject variability and device-specific noise. Efforts should explore FL for multimodal learning by integrating signals like EEG, ECG, and EMG and optimizing resource efficiency for wearable and edge devices.

6.4. Challenges in explainability

DL models are regarded as black boxes because their intricate mathematical operations are challenging for humans to interpret, leaving their decision-making processes unclear. This lack of transparency poses significant challenges, particularly in diagnosing diseases from biomedical signals, leading to skepticism among doctors about AI-driven decisions. As TL techniques become more complex, they reveal intricate interactions between domains that are challenging to interpret. A limited understanding of how target models use transferred knowledge and how it influences decisions can increase the risk of transfer failure and add extra time costs, hindering the broader adoption of these technologies. The interactions within transferred layers often need to be clarified and easier to unravel. Using explainable AI (XAI) methods can help us understand how TL-based diagnostic systems work, making it easier for doctors to trust these systems [225,226]. Keeping models interpretable is important for several reasons [227]. It helps us see how transferred knowledge affects learning in new areas, which allows for better fine-tuning and improvements. It also helps identify and fix errors, whether from biases in the original data or mistakes during the transfer process. Moreover, in fields like healthcare, it is crucial to make sure that decision-making processes are transparent, and justifiable. Without interpretability, complex TL methods might lead to unexpected problems, reduce trust, and perpetuate biases. Some researchers have utilized XAI techniques, including local interpretable model-agnostic explanations (LIME), shapley additive explanations (SHAP), and t-distributed stochastic neighbor embedding (t-SNE), alongside TL methods to analyze biomedical signals for different medical purposes. For future endeavors, employing different XAI techniques as a post-processing step in CAD systems based on TL models could yield valuable results in diagnosing diseases from biomedical signals.

6.5. Online transfer learning

A major limitation of current models is their reliance on static, pre-collected datasets, which limits their usefulness in real-world clinical settings that require real-time processing and learning from continuous biomedical data streams [228,229]. Online transfer learning (OTL) builds an online model for the target domain with knowledge drawn from source domains, thereby tackling the difficulties of predicting data that arrive sequentially when labeled data is limited. While OTL helps manage the issue of real-time data labeling, it still necessitates centralized access to data from both source and target domains, which could undermine data confidentiality and protection in the era of big data. In online mode, a major challenge arises from the significant differences between subjects and sessions in the data, which increases the risk of overfitting. In this scenario, the model demonstrates impressive performance during training but needs help to generalize effectively in practical testing conditions. This issue, compounded by the limited availability of information from the target domain, further complicates the model's ability to adjust to new and varying situations, thereby reducing its effectiveness in practical applications. A key application area in this new era is modeling real-time data, such as using wearable devices for disease detection, including epilepsy and heart monitoring. The OTL model integrates TL with online learning techniques, allowing for the continuous acquisition of training data rather than relying on pre-collected datasets, and has shown the highest accuracy in diagnostics. This capability supports adaptive knowledge transfer and facilitates timely updates to models, offering significant benefits for TL applications [74,230]. Future work should focus on advancing online TL to make it more effective in real-world applications. This involves creating methods to reduce differences between subjects and sessions, using DA strategies to limit overfitting, and finding ways to strengthen model robustness with limited target-domain data. Additionally, incorporating adaptive mechanisms that can continuously learn from new data in real- time could further boost the model's effectiveness and reliability in changing, practical settings.

6.6. Hardware resources

In the medical field, large pre-trained models have improved diagnostic accuracy and efficiency by analyzing patient data more effectively. However, as model sizes have grown significantly, fine-tuning the entire parameter set of pre-trained models has become increasingly expensive [231]. Although graphics processing units (GPUs) are typically employed for training DL networks, they are costly and not always easily available [232]. Also, in medical data analysis, high-resolution inputs are crucial for revealing subtle diagnostic details, but this significantly increases memory consumption during training. To address hardware challenges TL, future efforts can focus on a few practical

strategies. First, model compression techniques like weight pruning in filters, channels, or layers [233,234], network quantization [235,236], and knowledge distillation [237,238] can make TL models smaller and less demanding, making them more usable on devices with limited resources. Using efficient model designs, such as lightweight networks, can also help achieve robust performance with fewer resources. Edge computing and on-device learning allow models to run locally on devices, reducing the need for powerful servers [239,240]. Mixed-precision training, which adjusts the level of detail used in calculations based on available hardware, can further reduce memory and computation requirements [241]. Specialized, energy-efficient hardware like field programmable gate arrays (FPGAs) or application-specific integrated circuits (ASICs) can offer better processing with less power. Meanwhile, progressive training methods gradually increase model complexity to lower early resource demands. Future studies should concentrate on enhancing model efficiency using compression techniques, lightweight architectures, and energy-efficient hardware to address resource challenges in medical transfer learning applications.

6.7. Transformers

TL architectures, pre-trained on the ImageNet dataset for applications such as medical image classification, face challenges when applied directly to biomedical signals. Consequently, biomedical signals are typically transformed into 2D images using various techniques, such as STFT, before being input into TL architectures. This process may result in the loss of some properties of the signals. Recently, transformer architectures based on attention mechanisms have been proposed to directly analyze time-series data, particularly biomedical signals. Transformer architectures consist of two parts: the decoder and the encoder. They demonstrate high efficiency and performance on biomedical signal datasets with limited subjects, unlike TL networks. These architectures concentrate on specific sections of the data that contain the most important information. For future work, the latest pre-trained transformer models, such as BeIT, CoAT, DeIT, Visformer, and ViT networks, can be utilized on biomedical signals for various purposes. Generative pre-trained transformers (GPT) also offer great potential for research direction in biomedical signals, with potential for disease diagnosis and BCIs.

6.8. Self-supervised learning

To address the issue of insufficient labeled data, self-supervised learning (SSL) has shown to be a valuable alternative. SSL utilizes pretext tasks that eliminate the need for manual labeling, allowing models to learn useful features from significant volumes of unlabeled data [242]. These representations could subsequently be fine-tuned for specific tasks, thereby decreasing the dependency on large, labeled datasets. Recent applications of SSL in medicine have shown noteworthy progress, including advancements in disease diagnosis, enhanced model robustness, and improved handling of data noise [153,243]. Despite its advantages, SSL has limitations, such as the need for carefully designed pretext tasks to achieve useful representations. In the medical field, SSL has proven effective, with research indicating that SSL models can match or outperform the benchmark models while utilizing just 1 % of the labeled data required by conventional methods [244]. This helps address data scarcity and reduces the labor-intensive process of annotating medical data, which requires professional expertise. SSL's ability to handle multi-modal data such as images, videos, audio, and biomedical signals further accelerates model training. Pretext tasks in SSL are generally categorized into contrastive, predictive, and generative learning. For instance, Fedorov et al. [243] demonstrated that contrastive SSL techniques improved model performance and generalizability in Alzheimer's disease classification through MRI and fMRI data. Overall, SSL shows considerable promise in tackling the issue of limited labeled data in medical research and warrants further investigation.

6.9. Uncertainty quantification

Uncertainty quantification (UQ) in healthcare is essential for understanding and managing the variability in medical predictions, diagnoses, and outcomes of treatment. With the rise of AI-based models, accurately assessing uncertainty is crucial for safe decision-making. While XAI techniques enhance model interpretability, they do not tackle practical issues like decision reliability or capture overconfident predictions and vulnerability to adversarial attacks [245,246]. Effective uncertainty estimation helps manage factors such as incomplete data, noise, and inherent randomness, ensuring the reliability and safety of AI models. Biomedical signals are susceptible to artifacts that can unpredictably disrupt DL predictions. Therefore, UQ is particularly beneficial. It aids in detecting corrupted data and avoiding erroneous classifications. Moreover, since interpreting these signals demands considerable time and expertise, UQ helps determine which samples require expert review and which can be reliably classified automatically with a low probability of error. Future research in UQ holds promise for advancing the reliability of AI model predictions in real-world applications.

6.10. Few-shot TL

Few-shot learning techniques are designed to enable models to rapidly learn new tasks with minimal data. This capability is particularly crucial for applications involving biomedical signals, where the collection of a large volume of labeled data is commonly challenging, costly, and labor-intensive, and only a limited amount of data is typically available [247]. The few-shot learning task is a useful benchmark for evaluating ML models. However, the typical approach relies on the assumption that all classes have an equal number of data points, which is often impractical in real-world scenarios, where sample sizes can vary and imbalances, such as class or task imbalance, are common [248]. Future research should build on these insights to develop new few-shot learning methods that can maintain stable performance even in the presence of these realistic class imbalances. Additionally, most few-shot learning studies focus on unimodal settings, limiting their applicability in real-world scenarios where data are multimodal. In medicine, physicians integrate multimodal data for more accurate diagnoses. Thus, applying few-shot learning to multimodal contexts could improve model generalization and expand its practical use in medical applications.

7. Conclusion

TL has emerged as a vital approach in biomedical signals, effectively tackling the ongoing issue of insufficient data in real-world environments where mismatches exist in the feature or label distributions across source and target domains. This systematic review presents a comprehensive analysis of 239 studies employing TL for various biomedical signals over the past decade, categorizing them into label-based, feature-based, and learning-based methods. By systematically analyzing a wide variety of biosignals and their applications in disease detection, prediction, and BCI systems, this review highlights the growing use of TL methodologies across various domains. Most TL applications focus on EEG signals, particularly in BCIs and epilepsy detection and prediction. Our in-depth analysis of the studies, combined with a discussion of existing challenges, provides valuable insights to guide future research. Expanding access to high-quality, open biomedical signal datasets and addressing issues of domain mismatch, hardware resources, multimodality and real-time adaptability will be fundamental to the advancement of the field. By addressing the current limitations and focusing on innovative approaches, TL can unlock new possibilities for personalized medicine, efficient healthcare delivery, and the advancement of BCI technologies. This study is a valuable resource for

researchers aiming to develop more robust and effective TL techniques with improved generalizability for real-world applications. The comprehensive analysis and key insights presented in this review are expected to foster continued research and innovation in TL, opening the door for its wider application and greater impact across various real-world scenarios.

CRediT authorship contribution statement

Mahboobeh Jafari: Writing – review & editing, Writing – original draft, Visualization, Validation, Resources, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization. **Xiaohui Tao:** Writing – review & editing, Supervision. **Prabal Barua:** Writing – review & editing, Supervision. **Ru-San Tan:** Writing – review & editing. **U.Rajendra Acharya:** Writing – review & editing, Supervision, Investigation, Conceptualization.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

Mahboobeh Jafari reports financial support was provided by University of Southern Queensland. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgment

The first author has received the University of Southern Queensland (UniSQ) International Stipend and Fee Research Scholarship (DPHD), which is gratefully acknowledged.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.inffus.2025.102982.

Appendix A. Summary of TL studies on biomedical signals Incorporated into the review

Refer to Tables A.1, A.2, A.3, A.4, A.5, A.6, A.7, A.8, A.9, A.10, A.11, A.12, A.13, A.14, A.15, A.16, A.17, A.18, A.19

Table A.1
Summary of studies conducted on EEG-based transfer learning for epilepsy/ seizure analysis.

Ref.	Year	Dataset		TL model	Application	Results (%)
		Source	**Target**			
[249]	2022	CHB-MIT, TUSZ	CHB-MITCHB-MIT	EEGWaveNet	Seizure Detection	Acc = 98.26 Sen = 98.07 Spe = 98.45
[250]	2021	CHB-MIT	CHB-MIT	**Siamese Network** and CNN with MTL	Seizure Prediction	Pre-ictal-interictal Acc = 91.4 Sen = 92.45 Spe = 89.94
[110]	2021	ImageNet	Private	Pre-trained Alexnet, ResNet, Googlenet, **Densenet)**	Neonatal Seizure Detection	Acc = 99.17
[251]	2021	ImageNet	**CHB-MIT**, iNeuro	Pre-trained VGG-19, AlexNet, Inception-ResNet-V2,Inception-V3, ResNet-152	Epileptic State Detection	Acc = 98.97
[252]	2021	ImageNet	Real and Synthesized samples of Epilepsyecosy and CHB-MIT	Pre-trained VGG16, VGG19, ResNet50, **Inceptionv3**	Epileptic Seizure Prediction	Acc = 90.03
[109]	2022	ImageNet	Bonn University	Pre-trained VGG16, VGG19, ResNet50, InceptionV3, DenseNet121, Xception, NASNet, **ONASNet)**	Epileptic Seizure Detection	Acc = 99.67
[253]	2019	CHB-MIT	CHB-MIT	DCAE based SS	Epileptic Seizure Prediction	Acc = 99.6 Sen = 99.72 Spe = 99.6
[254]	2023	ImageNet	Bonn University	Pre-trained ResNet-50	Epileptic Signals Detection	Best Result for SVM Acc = 96.04 Sen = 95.34 Spe = 96.23 F1-Score = 95.78
[255]	2022	ImageNet	Bern-Barcelona	Pre-trained AlexNet, **Inception-V3**, Inception-ResNet-V2, ResNet-50, VGG-16	Focal and Non-focal Epileptic Seizure Detection	Acc = 92.27 Spe = 92.93 Pre = 92.84 Rec = 91.6 F1-Score = 92.21
[195]	2022	ImageNet	CHB-MIT, iNeuro	Ensemble of Pre-trained AlexNet, VGG-19, Inception-V3, ResNet-152, Inception-ResNet-V2	Epileptic Detection	Acc = 96.97
[256]	2022	CHB-MIT	CHB-MIT	DA (adversarial learning and Riemannian manifold)	Epileptic Detection and Prediction	Sen = 86.4 Detection, 82.2 Prediction
[257]	2016	University of Bonn	University of Bonn (same and different distributions)	Transductive TL FSL	Epilepsy Detection	Acc = 99.3 (Binary Classification)

(*continued on next page*)

Table A.1 (*continued*)

Ref.	Year	Dataset		TL model	Application	Results (%)
		Source	**Target**			
[258]	2017	University of Bonn	University of Bonn (same and different distributions)	Transductive TL TSK	Epileptic Seizures Detection	Acc > 95
[259]	2020	Private	Private	Fine-tuning CNN	Seizure Detection of Nocturnal Frontal Lobe Epilepsy	Acc > 94
[260]	2023	ImageNet	CHB-MIT, Siena	Pre-trained ResNet 50	Epilepsy Detection	Acc = 98.4 Sen = 97.92 FPR (/h) = 0.061
[261]	2019	Bonn	TUH	Knowledge and label space inductive TL	Epilepsy Detection	Acc = 81.21DWT Feature Extraction
[262]	2020	Bonn (different distributions)	Bonn (different distributions)	MST-TSK	Seizure Detection	Acc = 97.1
[263]	2023	ImageNet	CHB-MIT	Pre-trained Alexnet, Darknet19, Googlenet, Resnet50, Squeezenet	Epileptic Seizures Detection	Acc = 86.11 Pre = 84.21 Rec = 88.88
[264]	2023	Private	Private	DA (JPDDA)	Seizure Detection	Acc = 94.08
[265]	2023	CHB-MIT	CHB-MIT	Fine-tuning Hybrid Transformer based on inter-ictal, Both pre-ictal and inter-ictal	Epilepsy prediction	Sen = 91.7
[214]	2023	**CHB-MIT,** Kaggle	**CHB-MIT,** Kaggle	SSDA	Seizure Prediction	Sen = 88.8 AUC = 84.9
[114]	2018	Bonn, CHB-MIT (different and same distributions)	Bonn, **CHB-MIT** (different and same distributions)	Transductive TL	Epilepsy Detection	Acc = 94 Sen = 91.9 Spe = 93.2
[113]	2019	Bonn	Bonn	Transductive TL	Epilepsy Detection	Acc = 95.17
[266]	2020	Bonn (different distributions)	Bonn (different distributions)	Inductive TL	Epilepsy Detection	Acc = 93
[267]	2024	ImageNet	CHB-MIT	Pre-trained ResNet50	Epileptic Seizure Detection	Acc = 95.23 Sen = 99.54 Spe = 90.28
[112]	2014	Bonn (different and same distributions)	Bonn (different and same distributions)	TDAL	Epilepsy Detection	Acc > 93
[268]	2021	Private iEEG	Private iEEG	Pretrained LSTM	Wearable Seizure Detection	AUC = 82 Sen = 47 FAR = 7.2events/day
[269]	2023	CHB-MIT, Kaggle, Private	CHB-MIT, Kaggle, Private	Unsupervised TL	Epilepsy Detection	F1-Score = 81
[270]	2024	CHB-MIT	CHB-MIT	Multiview DA	Epilepsy Detection	Acc = 92.95 Sen = 88.13 Spe = 95.36 F1-Score = 89.28
[271]	2024	CHB-MIT	CHB-MIT	DA Based on pre-trained transformer	Children Seizure Prediction	Sen = 79.5 AUC = 81.4
[272]	2024	CHB-MIT, Private	Online	Fine-tuning pre-trained multihead model	Online Seizure Prediction	Acc = 64.9 Sen = 54.4 Spec = 66 AUC = 65.1
[273]	2024	ImageNet	Two Private datasets	Pre-trained MobileNetV2	Differentiating Epileptic and Psychogenic Non-Epileptic Seizures	Acc = 87.2 Sen = 87.9 Spe = 86.9
[274]	2024	ImageNet	Private	Pre-trained Inception, ResNet, DenseNet, VGG16, **VGG19**	Epileptiform abnormalities Detection	Acc = 77
[111]	2024	CHB-MIT, Clinical	**CHB-MIT**, Private	DA	Seizure Prediction	AUC = 95.4
[275]	2024	Private EPILEPSIAE	Private	Unsupervised TL (DCAE)	Seizure Prediction	FPR/h = 0.35

Table A.2
Summary of studies conducted on EEG-based transfer learning for schizophrenia detection.

Ref.	Year	Dataset		TL model	Application	Results (%)
		Source	**Target**			
[115]	2022	ImageNet	RepOD	Pre-trained VGG-16, ResNet50V2, InceptionV3, DenseNet121, **EfficientNetB0**	SZ Detection	Acc = 99.9 Sen = 99.54 Spe = 100 AUC = 99.8 F1-Score = 99.93
[276]	2021	ImageNet	Kaggle	Pre-trained **AlexNet**, VGG16, ResNet50	SZ Detection	Acc = 93.33 Spe = 91.07 Sen = 94.89 F1-Score = 94.4

(*continued on next page*)

Table A.2 (*continued*)

Ref.	Year	Dataset		TL model	Application	Results (%)
		Source	**Target**			
[277]	2020	ImageNet	RepOD	Pre-trained **ResNet-18,** AlexNet, Inception-V3, VGG-19	SZ Detection	Acc = 98.6 Sen = 99.65 Spe = 96.92
[278]	2022	ImageNet	Two Private Datasets	Feature Extraction data set1, Fine-tuning on dataset 2 (VGG-16)	SZ Detection	Acc = 83.2
[279]	2022	ImageNet	Lomonosov Moscow State University	Pre-trained ResNet 50, **ResNet 101**	SZ Detection	Acc = 97.7 Sen = 97.8 Spe = 97.7 F1-Score = 97.6

Table A.3
Summary of studies conducted on EEG-based transfer learning for ASD detection.

Ref.	Year	Dataset		TL model	Application	Results (%)
		Source	**Target**			
[280]	2024	ImageNet	Private	Pre-trained **SqueezeNet,** AlexNet, ResNet18, GoogLeNet, MobileNetV2, ShuffleNet, EfficientNetb0	ASD Detection	Acc = 87.8 Spe = 95.9 Rec = 87.8 Pre = 87.9 F1-Score = 87.8
[116]	2021	ImageNet	Private	Ensemble of Pre-trained MobileNetV2, ShuffleNet, SqueezeNet	ASD Detection	Acc = 96.44 Sen = 97.79 Pre = 97.19 Spe = 93.16 F1-Score = 97.49
[281]	2024	ImageNet	Private	Pre-trained **DenseNet-121**, ResNet-101	ASD Detection	Acc = 89.19

Table A.4
Summary of studies conducted on EEG-based transfer learning for mental workload/fatigue detection.

Ref.	Year	Dataset		TL model	Application	Results (%)
		Source	**Target**			
[282]	2024	ImageNet	Northeastern University (NEU)	Pre-trained ResNet18, Densenet 201, Shufflenet	Unfavorable Driving State Detection	Acc = 90 Pre = 89.51 Sen = 90.63 Spe = 89.38 F1-score = 90.06
[118]	2020	Private	Private	DA	Fatigue Detection	Pre = 88.02 Rec = 89.08 F1-Score = 88.46
[283]	2021	Private	Private	DA	Fatigue Mental State Prediction	ACC = 81.82 Rec = 81.76 Pre = 98.27 F1-Score = 81.98
[117]	2024	Sleep-EDF	Driving datasets	DA	Driving Fatigue Detection	Acc = 89.56 Pre = 85.58 Spe = 89.58 Sen = 89.17 F1-Score = 83.16
[284]	2024	Private	Private	DTDDAN	Driving Fatigue Detection	Acc = 92.2 Spe = 93.8 Sen = 89.8
[119]	2023	Private (1 task)	Private (another task)	Feature TL	MW Detection	Acc = 81.3 Sen = 81.2 Spe = 80.6
[285]	2023	Private	Private	Transductive TL	Driving Fatigue Detection	Acc = 88.26 Sen = 86.07 Spe = 91.37 Pre = 84.4 F1-Score = 84.07
[84]	2022	Private	Private	DA	MW Detection	Acc = 65.45

(*continued on next page*)

Table A.4 (*continued*)

Ref.	Year	Dataset		TL model	Application	Results (%)
		Source	**Target**			
[286]	2019	DEAP	Two Private datasets	DA (Pre-trained SAE)	Evaluating Operator MW	Acc = 90.71
[287]	2024	Private	Private	Pre-trained DS-CNN	MW Detection	Acc = 97.22
[288]	2024	Private (working memory)	Private (**English literature reading**), Driving fatigue	Pre-trained STTransformer	Mental Fatigue Detection	Acc = 89.66 Pre = 91.55 Rec = 89.17 F1-Score = 90.03

Table A.5
Summary of studies conducted on EEG-based transfer learning for emotion recognition.

Ref.	Year	Dataset		TL model	Application	Results (%)
		Source	**Target**			
[289]	2022	ImageNet	DEAP, MAHNOB-HCI, DREAMER	Pre-trained Xception, AlexNet, ShuffleNet, **ResNet-18**, DarkNet-19, Inception-V3	Emotion Recognition	MAHNOB-HCI: Acc = 95.25 F1-Score = 95.33 DREAMER: Acc = 96 F1-Score = 96.77 DEAP: Acc = 94.27 F1-Score = 94.74
[290]	2024	ImageNet	DEAP, MAHNOB-HCI	Ensemble of Pre-trained DenseNet-201, NasNet-Mobile, ResNet-50, EfficientNetB0, Inception-v3, Xception	Emotion Recognition	Acc = 87.56 Pre = 87.63 Sen = 87.48 F1-Score = 87.6
[121]	2023	Full channels datasets of CDEED, SEED-V, **SEED-IV**, DEAP	Few-channels datasets of DEAP, CDEED, **SEED-IV**, SEED-V	Fine-tuning GCN	Emotion Recognition	Acc = 89.41
[120]	2021	SEED (different subjects)	SEED (new subject)	DMATN	Emotion Recognition	Acc = 84.9
[291]	2024	DEAP	DREAMER, EEWD	Fine-tuning SS4-STANN	Emotion Recognition	Acc > 95
[292]	2024	**SEED,** DEAP, DREAMER (different subjects)	**SEED,** DEAP, DREAMER(new subject)	Adversarial DA	Emotion Recognition	Acc = 84.63
[293]	2024	SEED, SEED-IV	SEED, SEED-IV	FMLAN	Emotion Recognition	Acc = 90.96 (Cross-subject) Acc= 91.94 (Cross-session)
[294]	2022	**SEED**, DEAP(cross-dataset, **cross-subject**)	**SEED**, DEAP(cross-dataset, **cross-subject**)	MDTDDL	Emotion Recognition	Acc = 76.75 Pre = 77.51 Rec = 78.52
[85]	2023	SEED cross-subject)	SEED (cross-subject)	DA (domain selection)	Emotion Recognition	Acc = 70.3
[295]	2023	**SEED**, DEAP(different subjects)	**SEED**, DEAP(new subject)	MSDA	Emotion Recognition	Acc = 85.605
[66]	2023	**SEED,** DEAP, SEED-IV (different subjects, sessions)	**SEED,** DEAP, SEED-IV (new subject and new session)	MSDA	Emotion Recognition	Acc = 91.1 F1-Score = 91
[296]	2023	**SEED,** SEED-IV	**SEED,** SEED-IV	DA	Emotion Recognition	Acc = 93.6
[297]	2022	DEAP, **SEED**	DEAP, **SEED**	MTL-MSRN	Emotion Recognition	Acc = 87.05
[298]	2024	SEED, **SEED-IV** (different subjects)	SEED, **SEED-IV** (one subject)	DA	Emotion Recognition	Acc = 80.17
[299]	2021	ImageNet	Private	Pre-trained AlexNet, **VGG16**	Emotion Recognition	Acc = 73.28

Table A.6
Summary of studies conducted on EEG-based transfer learning for response prediction.

Ref.	Year	Dataset		TL model	Application	Results (%)
		Source	**Target**			
[300]	2023	ImageNet	TDBRAIN	Pre-trained VGG16, **Xception** EfcientNetB0	Response Prediction to rTMS in MDD	Xception -Bi-LSTM-Attention: Acc = 98.86

(*continued on next page*)

Table A.6 (*continued*)

Ref.	Year	Dataset		TL model	Application	Results (%)
		Source	**Target**			
						Sen = 80.2 Spe = 97.73 AUC = 83
[301]	2023	ImageNet	DBRAIN	Pre-trained VGG16, **EfficientNetB0**	Response Prediction to rTMS in MDD	Acc = 82.3 Sen = 80.2 Spe = 81.9 AUC = 83
[124]	2021	ImageNet	Hospital Universiti Sains Malaysia (HUSM)	Ensemble of pre-trained VGG16, Xception, DenseNet121, MobileNetV2, InceptionResNetV2	Response Prediction of drug in MDD	Acc = 96.5 Spe = 96.95 Sen = 96.01
[122]	2023	ImageNet	Hospital Universiti Sains Malaysia (HUSM)	Ensemble of pre-trained VGG16, Xception, Densenet121	Prediction of Treatment Outcome in MDD	Acc = 98.84 Sen = 98.80 Spe = 99.60
[123]	2023	ImageNet	Private	Ensemble of pre-trained VGG16, Xception, InceptionResNetV2, DenseNet121, EfficientNetB0	Prediction of the rTMS Treatment of MDD	Acc = 92.28 Sen = 98.56 Spe = 86 F1-Score = 92.75

Table A.7
Summary of studies conducted on EEG-based transfer learning for sleep stage applications.

Ref.	Year	Dataset		TL model	Application	Results (%)
		Source	**Target**			
[302]	2020	ImageNet	Physionet Sleep EDFx	Pre-trained Squeezenet	Sleep Stage Classification	Acc = 84.55 Sen = 77.06 Spe = 95.78
[303]	2022	ImageNet	Two Clinical datasets	Pre-trained **GoogLeNet**, CNN	Sleep Stage Classification	Acc = 91.2 Sen = 77 Pre = 75.9 Spe = 94.1
[304]	2022	MASS	Sleep-EDF	Pre-trained SeqSleepnet and DeepSleepnet	Sleep Stage Classification	Acc = 87.84
[125]	2023	EDF-set, SH-set, PhysioNet-100	EDF-set, SH-set, PhysioNet-100 (Cross-channel, Cross-subject, Cross-dataset)	DA	Sleep Stage Classification	Acc = 80.1
[238]	2024	MASS	Sleep-EDF	KD	Sleep Stage Classification	Acc = 86.5
[305]	2024	MASS2	MODA	Pre-trained (Convolutional Multi-dilated Block)	SS and KCs Detection	F1-Score = 80.5 SS, 83.7 KC

Table A.8
Summary of studies conducted on EEG-based transfer learning for BCI applications.

Ref.	Year	Dataset		TL model	Application	Results (%)
		Source	**Target**			
[306]	2019	ImageNet	BCI competition IV 2b	Pre-trained VGG-16	MI EEG Signal Classification	Acc = 74.2
[220]	2024	MI4C, **NICU,** SEED	MI4C, **NICU,** SEED	DS3TL	Cross-Subject EEG Classification	Acc = 82.29
[127]	2024	BCI IV 2a,**2b,** OpenBMI	MI IV 2a,**2b,** OpenBMI	Dynamic DA	MI Decoding	Acc = 85.19
[126]	2024	**OpenBMI,** Private	**OpenBMI,** Private	Domain Adversarial TL	P300 Detection	Acc = 97.54
[307]	2024	BCI Competition IV 2a, 2b, **BCI competition III 4a**	BCI Competition IV 2a, 2b, **BCI competition III 4a**	Adaptation of spatial-temporal features with dual regularization	MI Classification	Acc = 85.29
[308]	2024	BCI Competition IV 2a, **Autocalibration and Recurrent Adaptation (ARA)**	BCI Competition IV 2a, **Autocalibration and Recurrent Adaptation**	Selective-MDA	MI Classification	Acc = 67.07
[128]	2024	BCI Competition IV, OpenBMI, **SMR**	BCI Competition IV, OpenBMI, SMR	DA (MACNet)	BCI Applications	Acc = 85.3
[78]	2024	BCI Competition IV-2b, **CBCIC**	Online signal	Online TL (Pretrained classifiers)	Online BCI Applications	Error Rate = 31.85
[309]	2024	OpenBMI	Clinical Stoke Dataset	DA	MI Classification	Acc = 71.15
[310]	2024	BCI Competition IV 2a, **2b,** HGD	BCI Competition IV 2a, **2b**, HGD	SSCL-CSD	MI Classification	Acc = 82.34
[135]	2024	GIST	OpenBMI	DA	MI Classification	Acc = 74.28
[67]	2024	BCI Competition IV **2a,** 2b	BCI Competition IV **2a**, 2b	DA	MI Classification	Acc = 86.25
[311]	2024	BCI Competition IV 2a, **2b**	BCI Competition IV 2a, **2b**	Siamese DA	MI Classification	Acc = 87.52
[312]	2023	BCI Competition IV 2a, HGD	2008 BCI Competition IV 2a, HGD	CCSM-FT	MI Decoding	Acc = 80.26
[313]	2021	BCI competition IV 2a	BCI competition IV 2a (one subject)	Fine-tuning (CNN-LSTM)	MI Decoding	Acc = 81

(*continued on next page*)

Table A.8 (*continued*)

Ref.	Year	Dataset		TL model	Application	Results (%)
		Source	**Target**			
[314]	2024	Tsinghua, PhysioNet	Tsinghua, PhysioNet (one subject)	Domain Rectified TL	RSVP Classification	Cross Subject Classification AUC = 95.81 (Tsinghua) AUC = 74.48 (PhysioNet)
[131]	2022	BCI competition IV-2a	BCI competition IV-2a	MSFFCNN	EEG MI Multiclassification	Acc = 94.06 Kappa = 0.88
[134]	2023	Cross-dataset (BCIIV 2a, BCIIV 2b, HGD, Clinical)	Cross-dataset (BCIIV 2a, BCIIV 2b, HGD, Clinical)	Fine-tuning (ShallowNet, **DeepNet**, EEGNet) Cross-Dataset TL Based on Multi-Task Learning	MI Classification	Acc = 76.61
[315]	2022	IIa, IIb BCI Competition IV, **High Gamma**	IIa, IIb BCI Competition IV, **High Gamma**	Adversarial DA	MI Classification	Acc = 93.97
[316]	2024	**BCI IV 2a**, GigaDB	**BCI IV 2a**, GigaDB	DA	MI Classification	Acc = 81
[317]	2021	GIST, **KU**	GIST, **KU** (other subjects)	DA	BCI Applications	Acc = 73.32
[318]	2022	ImageNet	BCI Competition IV 2a	Pre-trained ResNet-50, **Inception-v3**	MI Classification	Acc = 92 Kappa = 88
[28]	2018	Physionet	Physionet	DA	MI-EEG Signal Classification	Acc = 86.49
[319]	2020	GIB-UVA ERP-BCI	GIB-UVA ERP-BCI (one subject)	Fine-tuning and Cross Subject TL (EEG-Inception)	Assistive ERP-based BCI	Acc = 84.6
[320]	2021	AudioSet	Tsinhgua BCI Lab	VGGish	SSVEP based BCI Classification	Acc = 82.2 F1-Score = 82.5
[321]	2020	BCI competition II, III	BCI competition II, III	Rule Adaptation (CNN)	P300 based BCI speller	P-Value = 0.00064
[322]	2021	EEG Motor Movement/Imagery	EEG Motor Movement/Imagery Dataset	Fine-tuning (1D- CNN)	MI Classification	Acc = 99.46
[323]	2022	BCI Competition 2008 2a MI-EEG	BCI Competition 2008 2a MI-EEG	Fine-tuning (CNN)	MI Classification	Acc = 83.65
[324]	2024	BCI Competition IV **2a**,2b	BCI Competition IV **2a**,2b (one subject)	MTLF	MI Classification	Acc = 73.69
[325]	2022	BCI Competition IV Dataset I, IIa, RSVP,ERN, MI BCI EEG and EMG	BCI Competition IV Dataset I, IIa, RSVP,ERN, MI BCI EEG and EMG (one subject)	Domain Selection	BCI Applications	Acc = 84.07
[326]	2022	BNCI Horizon 2020, **Private**	BNCI Horizon 2020, **Private**	Inter-task (VAE)	MI Classification	Acc = 83
[327]	2021	BCI Competition IV **1**, 2a	BCI Competition IV **1**,2a (one subject)	MFAR	MI Classification	Acc = 78.57
[328]	2023	ImageNet	Physionet, Private	Pre-trained SqueezeNet	MI Classification	Acc = 92.33
[329]	2024	BCI Competition IV, OpenBMI, **SMR**	BCI Competition IV, OpenBMI, **SMR**	Unsupervised DA	MI Classification	Acc = 85.83
[330]	2019	**BCI Competition III, Iva**, Private	**BCI Competition III, Iva,** Private	DTMKB	MI Classification	Acc = 87.6
[331]	2024	BCI Competition IV Datasets 2a,**2b** (cross subject and cross session)	BCI Competition IV Datasets 2a,**2b** (cross subject and cross session)	Fine-tuning (EISATC-Fusion)	MI Decoding	Acc = 71.23
[332]	2022	**BCI IV IIb**, BCI IV IIa, CLA MI	**BCI IV IIb**, BCI IV IIa, CLA MI (one subject)	JDA	MI Recognition	Acc = 76.65
[333]	2024	Two Privates	Two Privates	Subject-Adaptive TL	MI Classification	Acc = 88.70
[209]	2024	**OpenBMI,** BCICIV2a,ERN, RSVP	**OpenBMI,** BCICIV2a,ERN, RSVP	DA	BCI Applications	Acc = 74.68
[130]	2023	UCSD, Tsinghua, Private	UCSD, Tsinghua, Private	Domain Generalization	SSVEP-based BCI	Graphical results
[334]	2022	BCI-Horizon 2020	BCI-Horizon 2020	DA	MI Classification	Acc = 87
[335]	2022	ImageNet	BCI Competition IV-a, IV-b, BCI Competition III V, GigaDB	Ten Pre-trained Models	BCI Applications	ShuffleNet Acc = 99.52
[336]	2023	Online Clinical, SEED	Online Clinical, SEED	MSDA	aBCI	Acc = 72.17
[218]	2018	Private	Private	Inductive TL	BCI Applications	Acc = 68.89
[337]	2017	Private	Private	Online and Offline WAR	Reducing BCI Calibration Effort	BCA = 68.2
[338]	2024	BCI Competition IV 1, 2a	BCI Competition IV 1, 2a	DA (EA, TCA)	BCI Applications	MTS Acc = 80.9
[132]	2024	**BCI Competition IV 1,** IV-2a	**BCI Competition IV 1,** IV-2a	MSDA	MI Classification	Acc = 85.93
[133]	2024	BCI Competition III Iva, **BCI Competition IV IIb**	BCI Competition III Iva, **BCI Competition IV IIb**	SSMDA	MI Classification	Acc = 85.09
[129]	2024	MI2-2, MI2-4, MI2015	MI2-2, MI2-4, MI2015	Unsupervised Multisource-free DA	MI Classification	Acc = 69.44
[309]	2024	OpenBMI MI-EEG	Private (stroke)	DA	MI Detection in stroke patients	Healthy-to-stroke Acc = 71.15

Table A.9
Summary of studies on EEG-based transfer learning for other applications.

Ref.	Year	Dataset		TL model	Application	Results (%)
		Source	**Target**			
[339]	2021	ImageNet	MIT/BIH Polysomnographic	Pre-trained ResNet18, ResNet50, and ResNet101	Drowsiness Detection	Acc = 97.92
[340]	2019	ImageNet	Temple Hospital EEG Abnormal Corpus v2	Pre-trained AlexNet	Pathology Diagnosis	Acc = 89.13 Sen = 80.16 Spe = 96.67
[139]	2021	ImageNet	AEP	Pre-trained VGG16	Hearing Deficiency Diagnosis	Acc = 96.87
[341]	2023	ImageNet	Dreams K-complexes	Pre-trained **Darknet-53**, MobileNets, ResNet-18, SqueezeNet, Darknet-53-coco	Detection of K complexes	Pre = 99.44 Miss Rate = 54
[342]	2022	ImageNet	Dreams K-complexes database	Pre-trained AlexNet, ResNet-101, VGG19, **Inceptionv3**	Detection of K-complexes	Pre = 99.8
[343]	2022	Bonn, Institute of Psychiatry and Neurology SZ dataset	Bonn, Institute of Psychiatry and Neurology SZ dataset	TCA	EEG Classification	Acc > 95
[140]	2022	ImageNet	Private	Pre-trained **VGG16**, VGG19, ResNet50, InceptionV3	Alzheimer's Disease Diagnosis	Acc = 100 Pre = 100 Rec = 100 Spe = 100
[344]	2022	ImageNet	Bonn, IPiN, PhysioBank	Pre-trained AlexNet,LeNet, VGG19-Batch Normalization, **Inception-v4**, ResNeXt-50	Brain Diseases Detection (schizophrenia, Epilepcy, Sleep Disorder)	Sen = 97.46 Spe = 99.07 F1-Score = 97.42
[138]	2024	ImageNet	San Diego, Iowa	Pre-trained AlexNet	Parkinson's Disease Detection	Acc = 99.84
[345]	2024	BCI Competition IV 2a, **2b**, BNCI2015001	BCI Competition IV 2a, **2b,** BNCI2015001	DA (Parameter transfer and Riemannian space coordinate alignment)	Intention Recognition	Acc = 79.17
[346]	2024	**Bonn**, ISRUC	**Bonn**, ISRUC	KDTL (ConvNext (teacher) Mobilnet (student))	Mental Disorders Diagnosis	Acc = 97.28
[347]	2019	Clinical	Clinical (one subject)	Pre-trained CNN	Decoding of Covert Attention Focus	Acc = 90.7
[348]	2020	DEAP	DEAP	**Fine-tuning** and retraining (DNN)	Preference Prediction	Acc = 93 Pre = 93 Rec = 93
[136]	2023	ImageNet	Clinical	Pre-trained GoogLeNet, InceptionV3, RestNet50, RestNet101, EfcientNet01, **DenseNet201**	Authentication	Acc = 99.98
[349]	2022	Clinical	Clinical	Siamese Neural Network with Feature Extraction (VGG16)	Data Augmentation	Acc = 96
[137]	2021	MASS	Kermanshah University of Medical Sciences, Clinical	DA	Insomnia Detection	Acc = 90.9
[350]	2022	ImageNet	Clinical	Pre-trained Models	Delirium Detection	Acc = 97.17 (Alexnet)

Table A.10
Summary of studies conducted on ECG-based transfer learning for arrhythmia classification.

Ref.	Year	Dataset		TL model	Application	Results (%)
		Source	**Target**			
[141]	2024	ImageNet	MIT-BIH	Pre-trained MobileNet-V2	Arrhythmia Classification	Acc = 98.69 Pre = 95.8 Rec = 86.2 F1-Score = 90.8
[351]	2021	ImageNet	MIT-BIH	Pre-trained AlexNet, VGG-16, Inception-v3	Ventricular Arrhythmias Classification	Acc = 97.6 Sen = 98.2 Spe = 97.5 F1-Score = 97.9
[352]	2021	ImageNet	Clinical	Pre-trained Inception-V3	Arrythmia Classification	Acc = 98.46 Rec = 95.43 Spe = 96.75
[353]	2023	CPSC2018, Ningbo, PTB-XL	PTB-XL, CPSC2018, Ningbo	Pre-trained MAE	Arrythmia Classification	AUC = 77.1
[354]	2023	ImageNet	PhysioNet Computing Challenge 2017	Pre-trained **GoogLeNet,** ResNet-101, DenseNet-201, SqueezeNet	Driver Arrhythmia Classification	CWT Acc = 82.39 AUC = 96.5 F1-Score = 59.5

(*continued on next page*)

Table A.10 (*continued*)

Ref.	Year	Dataset		TL model	Application	Results (%)
		Source	**Target**			
[143]	2023	MIT-BIH, INCARTDB, SVDB, LTDB	MIT-BIH, INCARTDB, SVDB, LTDB (cross-domain)	Unsupervised DA	Arrythmia Classification	Acc = 94.4
[355]	2022	ImageNet	MIT-BIH arrhythmia, INCART, European ST-T MIT-BIH Supraventricular Arrhythmia	Pre-trained DenseNet	Arrhythmia Classification	Acc = 99.8 Pre = 98.34 Rec = 99.63 F1-Score = 98.91
[142]	2022	ImageNet	MIT-BIH NSR, BIDMC CHF MIT-BIH ARR,	Pre-trained AlexNet	Arrhythmia and Congestive Heart Failure	Acc = 99.06 Sen = 99.14 Pre = 99.32 Spe = 99.68
[356]	2022	ImageNet	Kaggle, MIT-BIH	Pre-trained ResNet50, **AlexNet**, SqueezeNet	Arrhythmia Classification	Acc = 98.38
[357]	2024	PTB-XL	CPSC2018	Fine-tuning ED-DGCN	Arrhythmia Classification	Pre = 84.9 Rec = 84.1 F1-score = 84.3

Table A.11
Summary of studies conducted on ECG-based transfer learning for atrial fibrillation detection.

Ref.	Year	Dataset		TL model	Application	Results (%)
		Source	**Target**			
[358]	2023	Phy2017	CPSC2018, MIT-BIH AF	Fine-tuning (MP-DLNet)	AF Detection	Acc = 85.7 F1-Score = 82
[144]	2023	MIT-BIH Arrhythmia, **AFDB**	MIT-BIH Arrhythmia, **AFDB**	DA	AF Detection	Acc = 99.13
[145]	2023	MIT-BIH AF	LT-AF	DA (Few-shot TL (Fine-tuning pre-trained Siamese NN))	AF Detection	Acc = 97.07
[359]	2019	Phy2017	Phy2017	FDResNets	AF Detection	AUC = 92.1 F1-Score = 89.9
[360]	2020	**AFDB,** Phy2017	**AFDB,** MITDB, Phy2017	DARN	AF Detection	Acc = 98.84 Sen = 98.97 Spe = 98.75
[146]	2021	Clinical	Clinical	Pre-trained CNN	AF Detection	AUC = 77.9 Sen = 100 Spe = 76
[361]	2024	ImageNet	PHYJ	Ensemble of Pre-trained VGG 19, AlexNet, Inception-ResNet-v2, ResNet152, Inceptionv3	Paroxysmal AF Detection	Acc = 90 Sen = 90 Spe = 90
[362]	2024	AFDB	MITDB, CPSC 2021	Pre-trained 1D-CNN	AF Detection	Series of 13 beats from AFDB Acc = 99.61 Sen = 99.44 Spe = 99.69
[363]	2021	ImageNet	Phy2017, PTB-XL, ICBEB2018	Pre-trained ResNet-18v2, **ResNet-34v2,** ResNet-50v2)	AF Detection	F1-score = 79.4
[147]	2024	AFDB, Phy2017, CPSC2018	AFDB, Phy2017, CPSC2018	DA and Pre-trained AlexNet, VGG11, **ResNet**	AF Detection	Phy2017→AFDB Acc = 97.58 F1-score = 96.83

Table A.12
Summary of studies conducted on ECG-based transfer learning for other cardiac applications.

Ref	Year	Dataset		TL model	Application	Result (%)
		Source	**Target**			
[364]	2022	ImageNet	PhysioNet 2016/CinC	Pre-trained ResNet	Cardiac Anomaly Detection	Acc = 89.55
[82]	2023	MIT-BIH	MIT-BIH	Pre-trained MadeGAN	Cardiac Detection	Acc = 95.5 Pre = 95.4 Rec = 95.6 F1-Scoe = 94.7
[148]	2023	ImageNet	Clinical	Pre-trained **Xception,** VGG-19, InceptionResNetV2, DenseNet-121	COPD Detection	Acc = 99.9 Sen = 99.6 Spe = 98.6 Pre = 98.3

(*continued on next page*)

Table A.12 (*continued*)

Ref	Year	Dataset		TL model	Application	Result (%)
		Source	**Target**			
						F1-Score = 98.9
[150]	2023	VFDB, SDDB, CUDB	Clinical	Pre-trained U-Net	Signal Restoration and Rhythm Analysis	Acc = 90.8 Spe = 90.6 Sen = 91.2
[149]	2023	ImageNet	12-lead ECG Mendeley	Pre-trained EfficientNetV2B2	MI Detection	Acc = 99.03 Spe = 99.49 Sen = 98.96 F1-Score = 99.01
[365]	2021	Clinical	Clinical	Pre-trained CNN	Rare Genetic Heart Disease Detection	Acc = 83 Sen = 80 Spe = 78
[366]	2022	ImageNet	MIT-BIH arrhythmia database	Pre-trained CNN	ECG Signals Detection	ACC = 99.13 Sen = 96.21 Pre = 96.55
[367]	2024	ImageNet	PTB	Pre-trained VGG-Net	MI Detection	Acc = 99.22 Spe = 99.49 Sen = 99.15

Table A.13
Summary of studies conducted on ECG-based transfer learning for other applications.

Ref.	Year	Dataset		TL model	Application	Results (%)
		Source	**Target**			
[156]	2021	ImageNet	PTB, CYBHI	Ensemble of Pre-trained ResNet and DenseNet	Human Recognition	Acc = 99.66
[157]	2019	ImageNet	(PTB)-ECG, (CU)-ECG	Pre-trained AlexNet, GoogLeNet, **ResNet**	Biometrics	Acc = 98.1 ((PTB)-ECG), 93.2 ((CU)-ECG)
[152]	2019	ImageNet	Clinical	Pre-trained AlexNet, VGGNet, ResNet, **DenseNet161**	DM Detection	Acc = 97.62 Sen = 100 Spe = 96.72 Pre = 92.33 F1-Score = 95.88
[153]	2022	AMIGOS, DREAMER,**WESAD**, SWELL	AMIGOS, DREAMER,**WESAD**, SWELL	Pre-trained CNN	Emotion Recognition	Acc = 96.9 F1-Score = 96.3
[154]	2024	ImageNet	ECG-ID, BIDMC, CINC-2011, CINC-2014, teleECG, MIT-BIH arrhythmia	Pre-trained AlexNet, VGG, ResNet	Noise Detection and Classification	Acc > 70 Sen > 70 Spe > 70 F1-Score > 70
[155]	2022	ImageNet	Clinical	Pre-trained VGG-19, EfficientNet-B4, DenseNet-121	COVID-19 Detection	Binary Class Acc = 100 Pre = 100 Rec = 100 F1-Score = 100
[368]	2018	ImageNet	PTB, CYBHI, LivDet2015, FVC 2004	Pre-trained VGG-Net	Biometric Authentication Systems	Acc = 98.97
[369]	2022	ImageNet	ECG-ID, **PTB**	Pre-trained VGG-16	Human Authentication	Acc = 99.39 Pre = 99.55 Rec = 99.49 F1-Score = 99.52
[151]	2022	Swell	WESAD	DA, Pre-trained 1D-VGG-16, CNN	Stress Prediction	Acc = 98.92 F1-Score = 99.3
[370]	2022	ImageNet	PhysioNet	Pre-trained **Xception**, ResNet-101, InceptionResNet-V2, DenseNet-201, Inception-V3, GoogLeNet, DarkNet-53	Driver's Stress Detection	Acc = 98.11 Sen = 98.33 Pre = 98.61 Spe = 98.89 F1-Score = 98.44

Table A.14
Summary of studies conducted on PPG-based transfer learning for blood pressure detection.

Ref.	Year	Dataset		TL model	Application	Results (%)
		Source	**Target**			
[371]	2024	Mindray	MIMIC	Pre-trained SEM-ResNet	Blood Pressure Estimation	MAE DBP = 0.01 SBP = 0.01
[372]	2023	Mindray	MIMIC	Pre-trained SE-ResNet, KD-Informer TL	Blood Pressure Estimation	MAE SBP = 0.02 DBP = 0.01
[373]	2022	MIMIC-III	MIMIC-III	Pre-trained BiLSTM-At	Blood Pressure Prediction	Acc = 92.2
[374]	2021	ImageNet	MIMIC-III(PPG), Clinical (rPPG)	Pre-trained AlexNet, **ResNet**	Blood Pressure Prediction	MAE SBP = 14.1 DBP = 11.2
[162]	2024	ImageNet	MIMIC II	Pre-trained VGG-16, ResNet-50, Xception, **DenseNet-121**, MobileNet-V2	Cuffless Blood Pressure Estimation	MAE MBP = 4.39 SBP = 5.63 DBP = 2.82
[164]	2024	MIMIC II, WESAD, PPG-DaLiA, Mindray	MIMIC II, WESAD, PPG-DaLiA, Mindray	Transformer-based self-supervised TL	Blood Pressure Estimation	MAE SBP = 0.85 DBP = 0.49
[375]	2022	ImageNet	MIMIC II	Pre-trained Inception v3, VGG-19, **AlexNet**	Blood Pressure Estimation	MAE SBP = 0.00 DBP = -0.04
[376]	2022	MIMIC III	New Subject	Pre-trained CNN-GRU	Blood Pressure Estimation	MAE SBP = 3.52 DBP = 2.20 mmHg
[163]	2021	MIMIC II	MIMIC II (New Subject)	Multi-domain Adversarial with DAE	Arterial Blood Pressure Prediction	MAE SBP = 5.424 DBP = 3.144 MBP = 2.885

Table A.15
Summary of studies conducted on PPG-based transfer learning for other applications.

Ref.	Year	Dataset		TL model	Application	Results (%)
		Source	**Target**			
[165]	2020	ImageNet	**MAHNOB-HCI**, ECGFitness, VIPL-HR, UBFC-RPPG	Pre-trained ResNet18	Heart Rate Estimation	MAE = 7.97
[166]	2023	DB_shape, DB_HT	DB_DT2	Pre-trained CNN	Type 2 Diabetes Detection	Different Results
[161]	2023	BIDMC, CapnoBase,	WESAD, **SensAI**	Pre-trained RRWaveNet	Respiratory Rate Estimation	MAE = 1.07 (Window = 64)
[377]	2020	Stanford University, IEEE Signal Processing Cup 2015	Private	Pre-trained CDAE	Cardiac Rhythm Detection	Sen = 98 Spe = 99 F1-Score = 93

Table A.16
Summary of studies conducted on PCG-based transfer learning for other heart diseases detection.

Ref.	Year	Dataset		TL model	Application	Results (%)
		Source	**Target**			
[169]	2022	ImageNet	Physionet challenge 2016, Yaseen GitHub database	Pre-trained **VGGNet-16**, ResNet-50	Heart Valve Disorders Detection	Resuts for FDPCT Acc = 92.19 Pre = 92.20 Rec = 92.20 Kappa = 98.91
[378]	2023	AudioSet-YouTube corpus	PhysioNet 2016, Yaseen GitHub database	Pre-trained YAMNet	Heart Valve Disease Classification	Acc = 99.83 Sen = 99.59 Spe = 99.9
[379]	2023	ImageNet	Clinical	Pre-trained **VGG-16**, VGG-19, Xception	MI Detection	Acc = 96.7 Sen = 97.4 Spe = 96
[380]	2022	ImageNet	PASCAL 2011	Pre-trained AlexNet	Heart Rate Detection	Acc > 95 Pre > 95 Rec > 95

(*continued on next page*)

Table A.16 (*continued*)

Ref.	Year	Dataset		TL model	Application	Results (%)
		Source	**Target**			
[381]	2022	ImageNet	Clinical	Pre-trained VGG16	CAD Detection	Acc = 91.19 Sen = 93.25 Spe = 89.13
[170]	2023	ImageNet	Hospital of Chongqing Medical University	Ensemble of VGG16, Xception, ResNet50, InceptionResNetv2	Diastolic Dysfunction Detection	Acc = 88.2 AUC = 91.1 Sen = 82.1 Spe = 92.7 F1-score = 89.2
[382]	2023	ImageNet	PASCAL, Kwon	Different Pre-trained Models	Cardiovascular Diseases Detection	Acc > 98 Sen > 98
[383]	2021	ImageNet	PhysioNet Challenge 2016	Pre-trained AlexNet	PCG Detection	Acc = 98
[384]	2022	ImageNet	PASCAL CHSC	Pre-trained VGG16, VGG19, **ResNet50**	Heartbeat Murmurs Detection	Acc = 87.65
[385]	2022	ImageNet	PhysioNet	Pre-trained AlexNet, VGG16, ResNet50, **InceptionV3**	Heart Sounds Detection	Acc = 88.06
[386]	2020	ImageNet	Pascal	Multiple Pre-trained Models	Cardiovascular Disease Detection	Acc = 89
[387]	2023	**PhysioNet/CinC Challenge 2016 heart sound**, Yaseen	**PhysioNet/CinC Challenge 2016 heart,** Pasc al HSDB, PASCAL	Pre-trained AmtNet	PCG signals Detection	Acc = 84.41 Rec = 74.86 Pre = 76.2 F1-Score = 75.49
[168]	2023	ImageNet	PhysioNet 2016	Ensemble of Pre-trained AlexNet, SqueezeNet, VGG19	Predicting Heart Sound	Acc = 99.2 Sen = 99.47 Spe = 99.09

Table A.17
Summary of studies conducted on multi-modalities-based transfer learning for different applications.

Ref.	Year	Modality	Dataset		TL model	Application	Results (%)
			Source	**Target**			
[388]	2021	ECG, PPG	MIT-BIH ECG	UMass Simband PPG	Pre-trained 1D-CNN	AF Detection	ECG Acc = 95.5 Sen = 94.50 Spe = 96 PPG Acc = 95.1 Sen = 94.6 Spe = 95.2
[177]	2024	ECG, HER, Genetic	Vanderbilt University Medical Center, Clinical, PTB, GEO	Vanderbilt University Medical Center, Clinical, PTB, GEO	ABCM TL	Cardiovascular disease Prediction	Acc = 93.5 Pre = 92 Rec = 94.5 AUC = 97.2
[178]	2020	EEG, EOG	Sleep-EDF, Sleep-EDFx	Sleep-EDF, Sleep-EDFx	Pre-trained CNN	Sleep Stage Detection	Acc = 93.58 Cohen's kappa = 0.899
[179]	2023	EEG, ECG, EMG	ImageNet	CAP Sleep Database v1.0.0 - PhysioNet	Pre-trained VGG16	Sleep Stage Detection	Acc = 95.43
[180]	2021	MEG, MRI	ImageNet	Clinical	Pre-trained AlexNet	Alzheimer's Disease Prediction	Acc = 93 sMCI versus pMCI (post conversion) Acc = 89 sMCI versus pMCI (pre-conversion) Acc = 74 HC versus sMCI vs pMCI
[389]	2024	ECG, WPS	MIT-BIH arrhythmia	Private	Pre-trained ConvNeXtTrans	Arrhythmia Detection	Acc = 99.26 Sen = 98.99 F1-Score = 99.07
[390]	2024	EEG, EOG	Sleep-EDF, MASS	Clinical (DOC: Disorders of consciousness patients)	Pre-trained CNN	Sleep Stage Detection	Acc = 81.1
[391]	2022	EEG, EMG	Clinical	Clinical	Pre-trained CNN	Gait Patterns Classification	Acc = 89.05
[173]	2020	EEG, EMG	4-class gesture classification (EMG),3-class mental state (EEG)	4-class gesture classification (EMG),3-class mental state (EEG)	Unsupervisd TL Cross-Domain	Biological Signal Processing	Acc = 93.82 (EMG-EEG)
[392]	2022	GSR, EEG, ECG/PPG,	ImageNet	DREAMER, AMIGOS, DEAP, MAHNOB-HCI	Pre-trained VGG-16	Affective Computing	Different Results
[181]	2022	EEG, fNIRS, pupil eye data (PE)	ImageNet	Clinical	Pre-trained AlexNet	Cognitive Workload Identification	Acc = 93

(*continued on next page*)

Table A.17 (*continued*)

Ref.	Year	Modality	Dataset		TL model	Application	Results (%)
			Source	**Target**			
[393]	2021	EEG, Human face video clips	DEAP, Video Data of Human Faces	DEAP, Video Data of Human Faces	Pre-trained 3D-CNN	Emotion Recognition	Acc = 96.13(Valence) Acc = 96.79 (Arousal)
[394]	2019	EEG, Facial Expression	Large face dataset	Small face dataset	Pre-trained CNN	Emotion Recognition	Acc = 68 Valence, 70 Arousal

Table A.18
Summary of studies on EMG-based transfer learning for gesture.

Ref.	Year	Dataset		TL model	Application	Results (%)
		Source	**Target**			
[185]	2019	ImageNet	Clinical	Pre-trained Inception-v3,VGG-16, VGG-19, **ResNet-50**, ResNet- 152,	n-EMG discharges Classification	Pre = 100 Rec = 100 F1-Score = 100
[188]	2021	NinaPro DB6	NinaPro DB6	DA (DAM)	Gesture Recognition	Acc = 97.22 (within-session)
[184]	2022	ImageNet	Clinical	Pre-trained VGG-16, VGG-19, AlexNet, SqueezeNet, GoogLeNet, **ResNet-50**, ResNet-34,	Hand gesture Recognition	Acc = 94.41
[186]	2021	ImageNet	NinaPro DB5	Pre-trained Alexnet	Wearable low-cost EMG Application in HCI	Acc = 70.4
[395]	2023	NinaPro DB1-E2, E3	Clinical	ISME+Pre-trained models	Gesture Recognition	Different results
[189]	2021	LongTermEMG, 3DC	LongTermEMG, 3DC	TADANN	Gesture Recognition	Online Acc = 57.14 Offline Acc = 88.44
[396]	2024	Hyser, 2 Clinicals	Hyser, 2 Clinicals (One Subject)	Unsupervised DA	Hand Gesture Recognition	Acc = 90
[340]	2020	NinaPro databases 2 and 3	NinaPro databases 2 and 3 (one trial and subject)	Pre-trained CNN	Hand movements Decoding	Different Results

Table A.19
Summary of studies on other modalities-based transfer learning.

Ref.	Year	Modality	Dataset		TL model	Application	Results (%)
			Source	**Target**			
[348]	2021	HR	Clinical	Clinical	DA (Kernel-based TL)	REM Sleep Stage Detection	Acc = 90.35 Sen = 83.74 Spe = 92.65
[191]	2022	BCG	ImageNet	Public BCG Signals	Pre-trained **AlexNet**, Res-Net50	Hypertension Detection	Acc = 96.87 (GT), 95.6 (STFT), 96.61 (SPWVD)
[192]	2023	fNIRS	ImageNet	Clinical	Pre-trained ResNet, VGG, MobileNet, EfficientNet, TinyNet	Walking Tasks	Acc = 81
[397]	2022	fNIRS	Technische Universität Berlin	Technische Universität Berlin	Pre-trained CNN	BCI	Acc = 96.5

Appendix B. Abbreviations

Refer to Table B.20.

Table B.20
Abbreviations.

A	Deep Multi-Source Adaptation Transfer Network (DMATN)
Accuracy (Acc)	Deep Neural Network (DNN)
Artificial Intelligence (AI)	Residual Network (ResNet)
Area Under Curve (AUC)	Deep Source Semi-Supervised Transfer Learning (DS3TL)
Atrial Fibrillation (AF)	Diabetes Mellitus (DM)
Attention-Based Cross-Modal (ABCM)	Diastolic Blood Pressure (DBP)
Attention Deficit Hyperactivity Disorder (ADHD)	Distribution Alignment Module (DAM)
Attentional Multi-Scale Temporal Network (AmtNet)	Domain Adaptation (DA)
Autism Spectrum Disorders (ASD)	Domain Adaptive Residual Network (DARN)
B	Domain Transfer Multiple Kernel Boosting (DTMKB)

(*continued on next page*)

Table B.20 (*continued*)

A	Deep Multi-Source Adaptation Transfer Network (DMATN)
Balanced Classification Accuracy (BCA)	Dynamic Threshold Distribution Domain Adaptation (DTDDAN)
Ballistocardiogram (BCG)	**E**
Blood Pressure (BP)	Electrocardiography (ECG)
Brain-Computer Interface (BCI)	Electroencephalogram (EEG)
C	Electromyography (EMG)
Cardiovascular Disease (CVD)	Electrooculogram (EOG)
China Physiological Signal Challenge 2018 Database (CPSC2018)	Encoder-Decoder Dynamic Graph Convolutional Network (ED-DGCN)
Chronic Obstructive Pulmonary Disease (COPD)	Euclidean alignment (EA)
Computer-Aided Diagnosis (CAD)	Explainable AI (XAI)
Convolutional Denoising Autoencoder (CDAE)	Eye Tracking (ET)
Convolutional Neural Network (CNN)	**F**
Coronary Artery Disease (CAD)	False Alarm Rate (FAR)
Cross-Channel Specific-Mutual Feature Transfer Learning (CCSM-FT)	False Negative (FN)
D	False Prediction Rate (FPR)
Deep Learning (DL)	False Positive (FP)
Fast Down-Sampling Residual Convolutional Neural Networks (FDResNets)	**P**
Fine-grained Mutual Learning Adaptation Network (FMLAN)	Paroxysmal Atrial Fibrillation (PAF)
Frequency-Domain Polynomial Chirplet Transform (FDPCT)	Photoplethysmography (PPG)
Functional Near-Infrared Spectroscopy (fNIRS)	PhysioNet/CinC Challenge Database 2017 (Phy2017)
Fuzzy Logic System (FSL)	Polysomnography (PSG)
G	Progressive Mild Cognitive Impairment (pMCI)
Galvanic Skin Response (GSR)	Precision (Prec)
Gated Recurrent Unit (GRU)	**Q**
Generative Adversarial Network (GAN)	**R**
Gradient Class Activation Mapping (Grad-CAM)	Recall (Rec)
Graph Convolutional Networks (GCN)	Recursive Feature Elimination (RFE)
H	Remote Photopletysmography (rPPG)
Heart Rate (HR)	Repetitive Transcranial Magnetic Stimulation (rTMS)
I	Respiration (RSP)
Inception Self-Attention Temporal Convolutional Network Fusion (EISATC-Fusion)	Respiratory Rate (RR)
Inception-MaxPooling-Squeeze- Excitation (IMSE)	Resting-State Functional Magnetic Resonance Imaging (rs-fMRI)
Intracranial Electroencephalography (iEEG)	**S**
J	Schizophrenia (SZ)
Joint distribution adaptation (JDA)	Selective Multisource Domain Adaptation (Selective-MDA)
Joint-Probability-Discrepancy-Based Domain Adaptation (JPDDA)	**T**
K	T2-Weighted (T2W)
K-complex (KC)	Takagi-Sugeno-Kang (TSK)
Knowledge Distillation (KD)	Task Functional Magnetic Resonance Imaging (T-fMRI)
Knowledge Distillation-based Transfer Learning (KDTL)	Transductive domain adaptive learning (TDAL)
L	Transferable Adaptive Domain Adversarial Neural Network (TADANN)
Lightweight Custom CNN (LC-CNN)	Transfer Learning (TL)
M	Transfer Component Analysis (TCA)
Machine Learning (ML)	True Negative (TN)
Magnetic Resonance Imaging (MRI)	True Positive (TP)
Magnetoencephalography (MEG)	**U**
Major Depressive Disorder (MDD)	**V**
Masked Autoencoder (MAE)	Variational Autoencoder (VAE)
Maximum Mean Discrepancy (MMD)	Visual Geometry Group (VGG)
Mean Absolute Error (MAE)	**W**
Mean Blood Pressure (MBP)	Wavelet Transform (WT)
Mean Square Error (MSE)	Weighted Adaptation Regularization With Source Domain Selection (wARSDS)
Mental Workload (MW)	**X**
Meta-Transfer Learning- Multi-Scale Residual Network (MTL-MSRN)	**Y**
MIT-BIH Arrhythmia Database (MITDB)	**Z**
MIT-BIH Atrial Fibrillation Database (AFDB)	
MIT-BIH Supraventricular Arrhythmia Database (SVBD)	
Motor Imagery (MI)	
Multi-Path Neural Network (MP-DLNet)	
Multi-Source Domain Adaptation (MSDA)	
Multi-Source Domain Transfer Learning Fusion (MTLF)	
Multi-Scale Feature Fused CNN (MSFFCNN)	
Multi-Source Fusion Adaptation regularization (MFAR)	
Multi-Source to Single-Target (MTS)	
Multiple-Source transfer learning-based Takagi–Sugeno–Kang fuzzy system (MST-TSK)	
Multi-Task Learning (MTL)	
Multiview Adversarial Contrastive Network (MACNet)	
Myocardial Infarction (MI)	
N	

Appendix C. Performance metrics

Refer to Table C.21.

Table C.21
Performance metrics.

Acc	$\boldsymbol{Acc} = \frac{\boldsymbol{TP}+\boldsymbol{TN}}{\boldsymbol{FP}+\boldsymbol{FN}+\boldsymbol{TP}+\boldsymbol{TN}}$
BCA [225]	$BCA = \frac{a_+ + a_-}{2}$
Rec	$Rec = \frac{TP}{FP+TP}$
Sen	$Sen = \frac{TP}{TP+FN}$
Spec	$Spec = \frac{TN}{FP+TN}$
Pre	$Pre = \frac{TP}{TP+FP}$
F1-Score	$F1-Score = \frac{2\ TP}{2\ TP+FP+FN}$
FPR	$FPR = \frac{FP}{Time}$
SBP [294]	$SBP = \left(\frac{1}{M}\right)\sum_{i=1}^{M} y_{p_i}^t$
DBP [294]	$DBP = \left(\frac{1}{M}\right)\sum_{i=1}^{M} y_{v_i}^t$
MBP [289]	$MBP = \frac{(2*DBP+SBP)}{3}$
MAE [300]	$MAE = \frac{1}{N}\sum_{i=1}^{N}\lvert p_i - \widehat{p}_i\rvert$
SD [300]	$SD = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(p_i-\widehat{p}_i)^2}$
Kappa [304]	$Kappa = \frac{c\ \times s - \sum_{l=1}^{5} p_l \times\ t_l}{s^2 - \sum_{l=1}^{5} p_l \times\ t_l}$

Appendix D. Public datasets

Refer to Table D.22.

Table D.22
Public datasets available related to this review.

Name of dataset	URL	Modality	Application	Description
CHB-MIT	https://physionet.org/content/chbmit/1.0.0/	EEG	Epileptic seizure	23 Patients, 198 Seizures, Sampling Frequency 256 (Hz)
Bonn	https://www.ukbonn.de/epileptologie/arbeitsgruppen/ag-lehnertz-neurophysik/	EEG	Epileptic seizure	22 Patients, Sampling Frequency 173.61 (Hz)
Bern-Barcelona	https://www.upf.edu/web/ntsa/downloads	EEG	Epileptic seizure	10 Patients, Sampling Frequency 512 (Hz)
RepOD	https://repod.icm.edu.pl/dataset.xhtml?persistentId=doi:10.18150/repod.0107441	EEG	Schizophrenia	14 patients, 14 normal, Sampling Frequency 256 (Hz)
Kaggle	https://www.kaggle.com/datasets/broach/button-tone-sz	EEG	Schizophrenia	23 patients, 22 normal, Sampling Frequency 1000 (Hz)
SEED	https://bcmi.sjtu.edu.cn/home/seed/seed.html	**EEG**, Peripheral physiological signals (Electromyography,Electro-Oculogram), Face video	Emotion recognition	15 (7 male, 8 female), Sampling Frequency 1000 (Hz)
SEED-IV	https://bcmi.sjtu.edu.cn/home/seed/seed-iv.html	**EEG**, Eye-tracking data	Emotion recognition	15 (7 male, 8 female), Sampling Frequency 1000 (Hz)
DEAP	https://www.eecs.qmul.ac.uk/mmv/datasets/deap/download.html	**EEG**, GSR, Respiration Amplitude, Skin Temperature, Blood Volume, Electromyogram, Electrooculogram	Emotion recognition	32 (16 male,16 female), Sampling Frequency 512 (Hz)
MANHOB-HCI	https://mahnob-db.eu/hci-tagging/	**EEG**, Visual, Audio, Eye Gaze, ECG, GSR, Respiration Amplitude, Skin temperature	Emotion recognition	27 (11 male, 16 female), Sampling Frequency 256 (Hz)
DREAMER	https://zenodo.org/records/546113	**EEG**, Peripheral physiological signals (Electrocardiogram)	Emotion recognition	25 (14 male,11 female), Sampling Frequency 128 (Hz)

(*continued on next page*)

Table D.22 (*continued*)

Name of dataset	URL	Modality	Application	Description
Sleep-EDF	https://www.physionet.org/content/sleep-edf/1.0.0/	PSG sleep recordings, containing EEG, EOG, chin EMG	Sleep	8 subjects, Sampling Frequency 100 (Hz)
Montreal Sleep Research (MASS)	https://borealisdata.ca/dataverse/MASS	PSG sleep recordings, containing EEG, EOG, ECG, EMG	Sleep	200 whole-night PSG recording, Sampling Frequency 256 (Hz)
TD-BRAIN	https://brainclinics.com/dua/	EEG	MDD	124 MDD patients, Sampling Frequency 500 (Hz)
BCI competition III a, b,iva,ivb,ivc	http://bbci.de/competition/iii/	EEG	BCI tasks	Different subjects, Different sampling frequencies (Hz)
BCI competition IV 2a,2b	https://www.bbci.de/competition/iv/	EEG	BCI tasks	9 subjects, Sampling Frequency 250 (Hz)
San Diego	https://openneuro.org/datasets/ds002778/versions/1.0.5	EEG	Parkinson's Disease	8 subjects, Sampling Frequency 256 (Hz)
MIT-BIH Arrhythmia Database	https://physionet.org/content/mitdb/1.0.0/	ECG	Arrhythmia	47 subjects, Sampling Frequency 360 (Hz)
AFDB	https://physionet.org/content/afdb/1.0.0/	ECG	Atrial Fibrillation	23 subjects, Sampling Frequency 250 (Hz)
SVBD	https://physionet.org/content/svdb/1.0.0/	ECG	Supraventricular arrhythmias	8 subjects, Sampling Frequency 360 (Hz)
PhysioNet/CinC Challenge 2017	https://physionet.org/content/challenge-2017/1.0.0/	ECG	Atrial fibrillation	516 subjects, Sampling Frequency 300 (Hz)
PTB Diagnostic ECG Database	https://www.physionet.org/content/ptbdb/1.0.0/	ECG	Different cardiac diseases	294 subjects, Sampling Frequency 1000 (Hz)
MIMIC III	https://physionet.org/content/mimiciii/1.4/	**PPG**, ECG, ABP, respiration,	Blood pressure	10,282 subjects, Sampling Frequency 125 (Hz)
PhysioNet 2016	https://www.physionet.org/challenge/2016/	PCG	Heart sound	3240 recording, Sampling Frequency 2000 (Hz)
PASCAL 2011	https://istethoscope.peterjbentley.com/heartchallenge/index.html	PCG	Heart sound	656 recordings for heart sound classification, 176 recordings for heart sound segmentation
NinaPro DB1	https://ninapro.hevs.ch/instructions/DB1.html	sEMG and kinematic data	Gesture Recognition	27 subjects

Data availability

No data was used for the research described in the article.

References

[1] O. Faust, M.G. Bairy, Nonlinear analysis of physiological signals: a review, J. Mech. Med. Biol. 12 (4) (2012), https://doi.org/10.1142/S0219519412400155.

[2] M. Jafari, D. Sadeghi, A. Shoeibi, H. Alinejad-Rokny, A. Beheshti, D.L. García, J. M. Gorriz, Empowering precision medicine: AI-driven schizophrenia diagnosis via EEG signals: A comprehensive review from 2002–2023, Applied Intelligence 54 (1) (2024) 35–79.

[3] D. Wang, Y. Shang, Modeling physiological data with deep belief networks, Int. J. Inf. Educ. Technol. 3 (5) (2013) 505–511, https://doi.org/10.7763/IJIET.2013.V3.326.

[4] W. Van Drongelen. Signal processing for neuroscientists, Academic press, 2018.

[5] S.R. Devasahayam, *Signals and systems in biomedical engineering: signal processing and physiological systems modeling*. 2012.

[6] J. Behar, A. Roebuck, M. Shahid, J. Daly, A. Hallack, N. Palmius, G.D. Clifford, SleepAp: an automated obstructive sleep apnoea screening application for smartphones, IEEE journal of biomedical and health informatics 19 (1) (2014) 325–331.

[7] D. Alvarez-Estevez, V. Moret-Bonillo, Computer-assisted diagnosis of the sleep apnea-hypopnea syndrome: a review, Sleep. Disord. 2015 (1) (2015) 237878, https://doi.org/10.1155/2015/237878.

[8] L.K. Morrell, F. Morrell, Evoked potentials and reaction times: a study of intra-individual variability, Electroencephalogr. Clin. Neurophysiol. 20 (6) (1966) 567–575, https://doi.org/10.1016/0013-4694(66)90021-6.

[9] B.J.A. Schijvenaars, *Intra-individual variability of the electrocardiogram: assessment and exploitation in computerized ECG analysis*. 2000.

[10] O. Faust, E.Y.K. Ng, Computer aided diagnosis for cardiovascular diseases based on ECG signals: a survey, J. Mech. Med. Biol. 16 (1) (2016), https://doi.org/10.1142/S0219519416400017.

[11] T.M. Alam, et al., A fuzzy inference-based decision support system for disease diagnosis, Comput. J. 66 (9) (2023) 2169–2180, https://doi.org/10.1093/COMJNL/BXAC068.

[12] M.R. Islam, et al., Emotion recognition from EEG signal focusing on deep learning and shallow learning techniques, IEEe Access. 9 (2021) 94601–94624, https://doi.org/10.1109/ACCESS.2021.3091487.

[13] B. Nakisa, M.N. Rastgoo, A. Rakotonirainy, F. Maire, V. Chandran, Automatic emotion recognition using temporal multimodal deep learning, IEEe Access. 8 (2020) 225463–225474, https://doi.org/10.1109/ACCESS.2020.3027026.

[14] R. Sharma, R.B. Pachori, P. Sircar, Automated emotion recognition based on higher order statistics and deep learning algorithm, Biomed. Signal. Process. Control 58 (2020) 101867, https://doi.org/10.1016/J.BSPC.2020.101867.

[15] I. Ullah, M. Hussain, E. ul H. Qazi, H. Aboalsamh, An automated system for epilepsy detection using EEG brain signals based on deep learning approach, Expert. Syst. Appl. 107 (2018) 61–71, https://doi.org/10.1016/J.ESWA.2018.04.021.

[16] M. Shamim Hossain, S.U. Amin, M. Alsulaiman, G. Muhammad, Applying deep learning for epilepsy seizure detection and brain mapping visualization, ACM Trans. Multimed. Comput. Commun. Appl. 15 (1s) (2019), https://doi.org/10.1145/3241056.

[17] S. Srinivasan, S. Dayalane, S. kumar Mathivanan, H. Rajadurai, P. Jayagopal, G. T. Dalu, Detection and classification of adult epilepsy using hybrid deep learning approach, Sci. Rep. 13 (1) (2023) 1–17, https://doi.org/10.1038/s41598-023-44763-7. *2023 13:1*.

[18] A.S. Heinsfeld, A.R. Franco, R.C. Craddock, A. Buchweitz, F. Meneguzzi, Identification of autism spectrum disorder using deep learning and the ABIDE dataset, Neuroimage Clin. 17 (2018) 16, https://doi.org/10.1016/J.NICL.2017.08.017.

[19] F. Ke, S. Choi, Y.H. Kang, K.A. Cheon, S.W. Lee, Exploring the structural and strategic bases of Autism spectrum disorders with deep learning, IEEe Access. 8 (2020) 153341–153352, https://doi.org/10.1109/ACCESS.2020.3016734.

[20] N.M. Rad, et al., Deep learning for automatic stereotypical motor movement detection using wearable sensors in autism spectrum disorders, Signal. Processing. 144 (2018) 180–191, https://doi.org/10.1016/J.SIGPRO.2017.10.011.

[21] Z. Mao, et al., Spatio-temporal deep learning method for ADHD fMRI classification, Inf. Sci. (N. Y.) 499 (2019) 1–11, https://doi.org/10.1016/J.INS.2019.05.043.

[22] H. Chen, Y. Song, X. Li, A deep learning framework for identifying children with ADHD using an EEG-based brain network, Neurocomputing. 356 (2019) 83–96, https://doi.org/10.1016/J.NEUCOM.2019.04.058.

[23] S. Saurabh, P.K. Gupta, Deep learning-based modified bidirectional LSTM network for classification of ADHD disorder, Arab. J. Sci. Eng. 49 (3) (2024) 3009–3026, https://doi.org/10.1007/S13369-023-07786-W/TABLES/8.

[24] L.L. Zeng, et al., Multi-site diagnostic classification of schizophrenia using discriminant deep learning with functional connectivity MRI, EBioMedicine 30 (2018) 74–85, https://doi.org/10.1016/J.EBIOM.2018.03.017.

[25] D. Ahmedt-Aristizabal, et al., Identification of children at risk of schizophrenia via deep learning and EEG responses, IEEe J. Biomed. Health Inform. 25 (1) (2021) 69–76, https://doi.org/10.1109/JBHI.2020.2984238.

[26] Z. Aslan, M. Akin, A deep learning approach in automated detection of schizophrenia using scalogram images of EEG signals, Phys. Eng. Sci. Med. 45 (1) (2022) 83–96, https://doi.org/10.1007/S13246-021-01083-2/TABLES/4.

[27] H. Hamid, N. Naseer, H. Nazeer, M.J. Khan, R. Khan, U.S. Khan, Analyzing classification performance of fNIRS-BCI for gait rehabilitation using deep neural networks, Sensors. (Basel) 22 (5) (2022), https://doi.org/10.3390/S22051932.

[28] H. Dose, J.S. Møller, H.K. Iversen, S. Puthusserypady, An end-to-end deep learning approach to MI-EEG signal classification for BCIs, Expert. Syst. Appl. 114 (2018) 532–542, https://doi.org/10.1016/J.ESWA.2018.08.031.

[29] P.J. Lin, et al., CNN-based prognosis of BCI rehabilitation using EEG from first session BCI training, IEEe Trans. Neural Syst. Rehabil. Eng. 29 (2021) 1936–1943, https://doi.org/10.1109/TNSRE.2021.3112167.

[30] P. Khan, et al., Machine learning and deep learning approaches for brain disease diagnosis: principles and recent advances, IEEe Access. 9 (2021) 37622–37655, https://doi.org/10.1109/ACCESS.2021.3062484.

[31] A.B.R. Shatte, D.M. Hutchinson, S.J. Teague, Machine learning in mental health: a scoping review of methods and applications, Psychol. Med. 49 (9) (2019) 1426–1448, https://doi.org/10.1017/S0033291719000151.

[32] L.G. J, C.F. van der Laak, Deep learning in histopathology: the path to the clinic, Nat. Med. 27 (5) (2021) 775–784.

[33] R. Aggarwal, et al., Diagnostic accuracy of deep learning in medical imaging: a systematic review and meta-analysis, Nat. Res. (2021), https://doi.org/10.1038/s41746-021-00438-z.

[34] K.D. Tzimourta, et al., Machine learning algorithms and statistical approaches for Alzheimer's disease analysis based on resting-State EEG recordings: a systematic review, Int. J. Neural Syst. 31 (5) (2021), https://doi.org/10.1142/S0129065721300023.

[35] K. Rasheed, et al., Machine learning for predicting epileptic seizures using EEG signals: a review, IEEe Rev. Biomed. Eng. 14 (2021) 139–155, https://doi.org/10.1109/RBME.2020.3008792.

[36] F. Zhuang, et al., A comprehensive survey on transfer learning, Proc. IEEE 109 (1) (2019) 43–76, https://doi.org/10.1109/JPROC.2020.3004555.

[37] Machine Learning And Deep Learning In Medical Data Analytics and Healthcare, ROUTLEDGE, 2022. Accessed: Aug. 16, 2024. [Online]. Available: https://www.routledge.com/Machine-Learning-and-Deep-Learning-in-Medical-Data-Analytics-and-Healthcare-Applications/Jena-Bhushan-Kose/p/book/9781032127644.

[38] J.E. van Engelen, H.H. Hoos, A survey on semi-supervised learning, Mach. Learn. 109 (2) (2020) 373–440, https://doi.org/10.1007/S10994-019-05855-6/FIGURES/5.

[39] S. Niu, Y. Liu, J. Wang, H. Song, A decade survey of transfer learning (2010–2020), IEEe Trans. Artif. Intell. 1 (2) (2020) 151–166, https://doi.org/10.1109/TAI.2021.3054609.

[40] A.M. Azab, H. Ahmadi, L. Mihaylova, M. Arvaneh, Dynamic time warping-based transfer learning for improving common spatial patterns in brain-computer interface, J. Neural Eng. 17 (1) (2020), https://doi.org/10.1088/1741-2552/AB64A0.

[41] H. Phan, O.Y. Chén, P. Koch, A. Mertins, M. De Vos, Deep transfer learning for single-channel automatic sleep staging with channel mismatch, in: European Signal Processing Conference 2019-September, 2019, https://doi.org/10.23919/EUSIPCO.2019.8902977.

[42] S. Atasever, N. Azginoglu, D.S. Terzi, R. Terzi, A comprehensive survey of deep learning research on medical image analysis with focus on transfer learning, Clin. Imaging 94 (2023) 18–41, https://doi.org/10.1016/J.CLINIMAG.2022.11.003.

[43] Y. LeCun, L. Bottou, Y. Bengio, P. Haffner, Gradient-based learning applied to document recognition, Proc. IEEE 86 (11) (1998) 2278–2323, https://doi.org/10.1109/5.726791.

[44] A. Krizhevsky, I. Sutskever, and G.E. Hinton, "ImageNet classification with deep convolutional neural networks", Accessed: Aug. 16, 2024. [Online]. Available: http://code.google.com/p/cuda-convnet/.

[45] H. Qassim, A. Verma, D. Feinzimer, Compressed residual-VGG16 CNN model for big data places image recognition, in: 2018 IEEE 8th Annual Computing and Communication Workshop and Conference, CCWC 2018 2018-January, 2018, pp. 169–175, https://doi.org/10.1109/CCWC.2018.8301729.

[46] K. He, X. Zhang, S. Ren, J. Sun, Deep residual learning for image recognition, in: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition 2016-December, 2015, pp. 770–778, https://doi.org/10.1109/CVPR.2016.90.

[47] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, Z. Wojna, Rethinking the inception architecture for computer vision, in: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition 2016-December, 2015, pp. 2818–2826, https://doi.org/10.1109/CVPR.2016.308.

[48] G. Huang, Z. Liu, L. Van Der Maaten, K.Q. Weinberger, Densely connected convolutional networks, in: *Proceedings - 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017* 2017-January, 2016, pp. 2261–2269, https://doi.org/10.1109/CVPR.2017.243.

[49] F. Chollet, Xception: deep learning with depthwise separable convolutions, in: *Proceedings - 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017* 2017-January, 2016, pp. 1800–1807, https://doi.org/10.1109/CVPR.2017.195.

[50] L. Ding, et al., Exploring the potential of pretrained CNNs and time-frequency methods for accurate epileptic EEG classification: a comparative study, Biomed. Phys. Eng. Express. 10 (4) (2024) 045023, https://doi.org/10.1088/2057-1976/AD3CDE.

[51] C. Szegedy, et al., Going deeper with convolutions, in: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition 07-12-June-2015, 2014, pp. 1–9, https://doi.org/10.1109/CVPR.2015.7298594.

[52] K. Weiss, T.M. Khoshgoftaar, D.D. Wang, A survey of transfer learning, J. Big. Data 3 (1) (2016) 1–40, https://doi.org/10.1186/S40537-016-0043-6/TABLES/6.

[53] S.J. Pan, Q. Yang, A survey on transfer learning, IEEe Trans. Knowl. Data Eng. 22 (10) (2010) 1345–1359, https://doi.org/10.1109/TKDE.2009.191.

[54] C. Tan, F. Sun, T. Kong, W. Zhang, C. Yang, C. Liu, A survey on deep transfer learning, in: Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics) 11141, LNCS, 2018, pp. 270–279, https://doi.org/10.1007/978-3-030-01424-7_27.

[55] Humpire-Mamani, G. E., Jacobs, C., Prokop, M., van Ginneken, B., & Lessmann, N. (2023). Transfer learning from a sparsely annotated dataset of 3D medical images. *arXiv preprint arXiv:2311.05032*.

[56] C.Y. Wee, C. Liu, A. Lee, J.S. Poh, H. Ji, A. Qiu, Cortical graph neural network for AD and MCI diagnosis and transfer learning across populations, Neuroimage Clin. 23 (2019), https://doi.org/10.1016/j.nicl.2019.101929.

[57] H. Choi, Y.K. Kim, E.J. Yoon, J.Y. Lee, D.S. Lee, Cognitive signature of brain FDG PET based on deep learning: domain transfer from Alzheimer's disease to Parkinson's disease, Eur. J. Nucl. Med. Mol. ImAging 47 (2) (2020) 403–412, https://doi.org/10.1007/s00259-019-04538-7.

[58] A. Hosna, E. Merry, J. Gyalmo, Z. Alom, Z. Aung, M.A. Azim, Transfer learning: a friendly introduction, J. Big. Data 9 (1) (2022), https://doi.org/10.1186/s40537-022-00652-w.

[59] W. Li, W. Huan, B. Hou, Y. Tian, Z. Zhang, A. Song, Can emotion Be transferred? - A review on transfer learning for EEG-based emotion recognition, IEEe Trans. Cogn. Dev. Syst. 14 (3) (2022) 833–846, https://doi.org/10.1109/TCDS.2021.3098842.

[60] Z. Wan, R. Yang, M. Huang, N. Zeng, X. Liu, A review on transfer learning in EEG signal analysis, Neurocomputing. 421 (2021) 1–14, https://doi.org/10.1016/J.NEUCOM.2020.09.017.

[61] R. Raina, A. Battle, H. Lee, B. Packer, A.Y. Ng, June). Self-taught learning: transfer learning from unlabeled data, in: *Proceedings of the 24th international conference on Machine learning*, 2007, pp. 759–766.

[62] Huang, J., Gretton, A., Borgwardt, K., Schölkopf, B., & Smola, A. (2006). Correcting sample selection bias by unlabeled data. *Advances in neural information processing systems, 19*.

[63] M. Sugiyama, T. Suzuki, S. Nakajima, H. Kashima, P. Von Bünau, M. Kawanabe, Direct importance estimation for covariate shift adaptation, Ann. Inst. Stat. Math. 60 (4) (2008) 699–746, https://doi.org/10.1007/s10463-008-0197-x.

[64] J. Lu, V. Behbood, P. Hao, H. Zuo, S. Xue, G. Zhang, Transfer learning using computational intelligence: a survey, Knowl. Based. Syst. 80 (2015) 14–23, https://doi.org/10.1016/J.KNOSYS.2015.01.010.

[65] Z. Lan, O. Sourina, L. Wang, R. Scherer, G.R. Muller-Putz, Domain adaptation techniques for EEG-based emotion recognition: a comparative study on two public datasets, IEEe Trans. Cogn. Dev. Syst. 11 (1) (2019) 85–94, https://doi.org/10.1109/TCDS.2018.2826840.

[66] Q. She, C. Zhang, F. Fang, Y. Ma, Y. Zhang, Multisource Associate domain adaptation for cross-subject and cross-session EEG emotion recognition, IEEe Trans. Instrum. Meas. 72 (2023), https://doi.org/10.1109/TIM.2023.3277985.

[67] W. Liu, C. Guo, C. Gao, A cross-session motor imagery classification method based on riemannian geometry and deep domain adaptation, Expert. Syst. Appl. 237 (2024), https://doi.org/10.1016/j.eswa.2023.121612.

[68] G. Wang, M. Chen, Z. Ding, J. Li, H. Yang, P. Zhang, Inter-patient ECG arrhythmia heartbeat classification based on unsupervised domain adaptation, Neurocomputing. 454 (2021) 339–349, https://doi.org/10.1016/J.NEUCOM.2021.04.104.

[69] Y. Ganin, V. Lempitsky, Unsupervised domain adaptation by backpropagation, in: 32nd International Conference on Machine Learning 2, ICML, 2015, pp. 1180–1189. Sep. 2014, Accessed: Aug. 16, 2024. [Online]. Available, https://arxiv.org/abs/1409.7495v2.

[70] H. Chang, J. Han, C. Zhong, A.M. Snijders, J.H. Mao, Unsupervised transfer learning via multi-scale convolutional sparse coding for biomedical applications, IEEe Trans. Pattern. Anal. Mach. Intell. 40 (5) (2018) 1182, https://doi.org/10.1109/TPAMI.2017.2656884.

[71] J.J. Bird, J. Kobylarz, D.R. Faria, A. Ekart, E.P. Ribeiro, Cross-domain MLP and CNN transfer learning for biological signal processing: EEG and EMG, IEEe Access. 8 (2020) 54789–54801, https://doi.org/10.1109/ACCESS.2020.2979074.

[72] K.S. Barsim, W. Zheng, B. Yang, Ensemble learning to EEG-based brain computer interfaces with applications on P300-Spellers, in: *Proceedings - 2018 IEEE International Conference on Systems, Man, and Cybernetics, SMC 2018*, 2018, pp. 631–638, https://doi.org/10.1109/SMC.2018.00116.

[73] R. Ramele, A.J. Villar, J.M. Santos, EEG waveform analysis of P300 ERP with applications to Brain computer interfaces, Brain Sci. 8 (11) (2018), https://doi.org/10.3390/BRAINSCI8110199.

[74] P. Zhao, S.C.H. Hoi, J. Wang, B. Li, Online transfer learning, Artif. Intell. 216 (2014) 76–102, https://doi.org/10.1016/J.ARTINT.2014.06.003.

[75] D. Wu, V.J. Lawhern, S. Gordon, B.J. Lance, C.T. Lin, Driver drowsiness estimation from EEG signals using online weighted adaptation regularization for regression (OwARR), IEEE Trans. Fuzzy Syst. 25 (6) (2017) 1522–1535, https://doi.org/10.1109/TFUZZ.2016.2633379.

[76] Y. Jiang, Y. Zhang, C. Lin, D. Wu, C.T. Lin, EEG-based driver drowsiness estimation using an online multi-view and transfer TSK fuzzy system, IEEE Trans. Intell. Transp. Syst. 22 (3) (2021) 1752–1764, https://doi.org/10.1109/TITS.2020.2973673.

[77] Y. Zhang, et al., Epilepsy signal recognition using online transfer TSK fuzzy classifier underlying classification error and joint distribution consensus regularization, IEEE/ACM. Trans. Comput. Biol. Bioinform. 18 (5) (2021) 1667–1678, https://doi.org/10.1109/TCBB.2020.3002562.

[78] H. Wu, et al., Online privacy-preserving EEG classification by source-free transfer learning, IEEE Trans. Neural Syst. Rehabil. Eng. (2024), https://doi.org/10.1109/TNSRE.2024.3445115.

[79] A.J. Casson, D.C. Yates, S.J.M. Smith, J.S. Duncan, E. Rodriguez-Villegas, Wearable electroencephalography, IEEE Eng. Med. Biol. Magazine 29 (3) (2010) 44–56, https://doi.org/10.1109/MEMB.2010.936545.

[80] J. Huang, A.J. Smola, A. Gretton, K.M. Borgwardt, B. Schölkopf, Correcting sample selection bias by unlabeled data, Adv. Neural Inf. Process. Syst. 19 (2006).

[81] Tan, B., Zhang, Y., Pan, S., & Yang, Q. (2017, February). Distant domain transfer learning. In *Proceedings of the AAAI conference on artificial intelligence* (Vol. 31, No. 1).

[82] Z. Wang, S. Stavrakis, B. Yao, Hierarchical deep learning with generative adversarial network for automatic cardiac diagnosis from ECG signals, Comput. Biol. Med. 155 (2023), https://doi.org/10.1016/j.compbiomed.2023.106641.

[83] P. Kora, et al., Transfer learning techniques for medical image analysis: a review, Biocybern. Biomed. Eng. 42 (1) (2022) 79–107, https://doi.org/10.1016/J.BBE.2021.11.004.

[84] Y. Zhou, et al., Cross-task cognitive workload recognition based on EEG and domain adaptation, IEEE Trans. Neural Syst. Rehabil. Eng. 30 (2022) 50–60, https://doi.org/10.1109/TNSRE.2022.3140456.

[85] Y. Ma, W. Zhao, M. Meng, Q. Zhang, Q. She, J. Zhang, Cross-subject emotion recognition based on domain similarity of EEG signal transfer learning, IEEE Trans. Neural Syst. Rehabil. Eng. 31 (2023) 936–943, https://doi.org/10.1109/TNSRE.2023.3236687.

[86] D. Moher, A. Liberati, J. Tetzlaff, D.G. Altman, Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement, Int. J. Surg. 8 (5) (2010) 336–341, https://doi.org/10.1016/J.IJSU.2010.02.007.

[87] Y. Han and C. Ding, "Foundation models in electrocardiogram: a review," Oct. 2024, [Online]. Available: http://arxiv.org/abs/2410.19877.

[88] M. Li, D. Xu, Transfer Learning in Motor Imagery Brain Computer Interface: A Review, Shanghai Jiaotong University, 2024, https://doi.org/10.1007/s12204-022-2488-4.

[89] B. Abibullaev, A. Keutayeva, A. Zollanvari, Deep learning in EEG-based BCIs: a comprehensive review of transformer models, advantages, challenges, and applications, IEEe Access. 11 (2023) 127271–127301, https://doi.org/10.1109/ACCESS.2023.3329678.

[90] D. Wu, Y. Xu, B.L. Lu, Transfer learning for EEG-based brain-computer interfaces: a review of progress made since 2016, IEEe Trans. Cogn. Dev. Syst. 14 (1) (2022) 4–19, https://doi.org/10.1109/TCDS.2020.3007453.

[91] M. Heenaye-Mamode Khan, P. Reesaul, M.M. Auzine, A. Taylor, Detection of Alzheimer's disease using pre-trained deep learning models through transfer learning: a review, Artif. Intell. Rev. 57 (10) (2024), https://doi.org/10.1007/s10462-024-10914-z.

[92] A. Apicella, et al., Toward Cross-Subject and Cross-Session Generalization in EEG-based Emotion recognition: Systematic Review, Taxonomy, and Methods, Elsevier B.V, 2024, https://doi.org/10.1016/j.neucom.2024.128354.

[93] A. Ray, M.H. Kolekar, Transfer learning and its extensive appositeness in human activity recognition: a survey, Expert. Syst. Appl. 240 (2024) 122538, https://doi.org/10.1016/J.ESWA.2023.122538.

[94] L. Chato, E. Regentova, Survey of Transfer Learning Approaches in the Machine Learning of Digital Health Sensing Data, Multidisciplinary Digital Publishing Institute (MDPI), 2023, https://doi.org/10.3390/jpm13121703.

[95] R. Supakar, P. Satvaya, P. Chakrabarti, A deep learning based model using RNN-LSTM for the detection of schizophrenia from EEG data, Comput. Biol. Med. 151 (2022), https://doi.org/10.1016/j.compbiomed.2022.106225.

[96] G. Sannino, G. De Pietro, A deep learning approach for ECG-based heartbeat classification for arrhythmia detection, Future Gener. Comput. Syst. 86 (Sep. 2018) 446–455, https://doi.org/10.1016/j.future.2018.03.057.

[97] I. Ullah, M. Hussain, H. Aboalsamh, An automated system for epilepsy detection using EEG brain signals based on deep learning approach, Expert Systems with Applications 107 (2018) 61–71.

[98] Z. Yin, J. Zhang, Cross-session classification of mental workload levels using EEG and an adaptive deep learning model, Biomed. Signal. Process. Control 33 (2017) 30–47, https://doi.org/10.1016/j.bspc.2016.11.013.

[99] S. Chambon, M.N. Galtier, P.J. Arnal, G. Wainrib, A. Gramfort, A deep learning architecture for temporal sleep stage classification using multivariate and multimodal time series, IEEE Trans. Neural Syst. Rehabil. Eng. 26 (4) (Apr. 2018) 758–769, https://doi.org/10.1109/TNSRE.2018.2813138.

[100] W. Liu, W.-L. Zheng, and B.-L. Lu, "Emotion recognition using Multimodal deep learning," 2016, pp. 521–529. doi: 10.1007/978-3-319-46672-9_58.

[101] M. Gao, J. Mao, A novel active rehabilitation model for stroke patients using electroencephalography signals and Deep learning technology, Front. Neurosci. 15 (2021), https://doi.org/10.3389/fnins.2021.780147.

[102] X. Bi, H. Wang, Early Alzheimer's disease diagnosis based on EEG spectral images using deep learning, Neural Netw. 114 (2019) 119–135, https://doi.org/10.1016/j.neunet.2019.02.005.

[103] C. da Silva Lourenço, M.C. Tjepkema-Cloostermans, M.J.A.M. van Putten, Efficient use of clinical EEG data for deep learning in epilepsy, Clin. Neurophysiol. 132 (6) (2021) 1234–1240, https://doi.org/10.1016/j.clinph.2021.01.035.

[104] D. Ahmedt-Aristizabal, et al., Identification of children at risk of schizophrenia via deep learning and EEG responses, IEEe J. Biomed. Health Inform. 25 (1) (2021) 69–76, https://doi.org/10.1109/JBHI.2020.2984238.

[105] A.M. Alvi, S. Siuly, H. Wang, K. Wang, F. Whittaker, A deep learning based framework for diagnosis of mild cognitive impairment, Knowl. Based. Syst. 248 (2022), https://doi.org/10.1016/j.knosys.2022.108815.

[106] P. Saha, M. Abdul-Mageed, and S. Fels, "SPEAK YOUR MIND! towards imagined speech recognition with hierarchical deep learning," Apr. 2019, [Online]. Available: http://arxiv.org/abs/1904.05746.

[107] M.R. Islam, et al., Emotion recognition from EEG signal focusing on deep learning and shallow learning techniques, IEEe Access. 9 (2021) 94601–94624, https://doi.org/10.1109/ACCESS.2021.3091487.

[108] N.E. Elsayed, A.S. Tolba, M.Z. Rashad, T. Belal, S. Sarhan, A deep learning approach for brain computer interaction-motor execution EEG signal classification, IEEe Access. 9 (2021) 101513–101529, https://doi.org/10.1109/ACCESS.2021.3097797.

[109] Z. Song, B. Deng, J. Wang, G. Yi, W. Yue, Epileptic seizure detection using brain-rhythmic recurrence biomarkers and ONASNet-based transfer learning, IEEE Trans. Neural Syst. Rehabil. Eng. 30 (2022) 979–989, https://doi.org/10.1109/TNSRE.2022.3165060.

[110] A. Caliskan, S. Rencuzogullari, Transfer learning to detect neonatal seizure from electroencephalography signals, Neural Comput. Appl. 33 (18) (Sep. 2021) 12087–12101, https://doi.org/10.1007/s00521-021-05878-y.

[111] W. Cui, et al., Deep multiview module adaption transfer network for subject-specific EEG recognition, IEEe Trans. Neural Netw. Learn. Syst. (2024), https://doi.org/10.1109/TNNLS.2024.3350085.

[112] C. Yang, Z. Deng, K.S. Choi, Y. Jiang, S. Wang, Transductive domain adaptive learning for epileptic electroencephalogram recognition, Artif. Intell. Med. 62 (3) (2014) 165–177, https://doi.org/10.1016/j.artmed.2014.10.002.

[113] L. Xie, Z. Deng, P. Xu, K.S. Choi, S. Wang, Generalized hidden-mapping transductive transfer learning for recognition of epileptic electroencephalogram signals, IEEe Trans. Cybern. 49 (6) (2019) 2200–2214, https://doi.org/10.1109/TCYB.2018.2821764.

[114] Z. Deng, P. Xu, L. Xie, K.S. Choi, S. Wang, Transductive joint-knowledge-transfer TSK FS for recognition of epileptic EEG signals, IEEE Trans. Neural Syst. Rehabil. Eng. 26 (8) (2018) 1481–1494, https://doi.org/10.1109/TNSRE.2018.2850308.

[115] S. Bagherzadeh, M.S. Shahabi, A. Shalbaf, Detection of schizophrenia using hybrid of deep learning and brain effective connectivity image from electroencephalogram signal, Comput. Biol. Med. 146 (2022), https://doi.org/10.1016/j.compbiomed.2022.105570.

[116] M. Baygin, et al., Automated ASD detection using hybrid deep lightweight features extracted from EEG signals, Comput. Biol. Med. 134 (2021), https://doi.org/10.1016/j.compbiomed.2021.104548.

[117] F. Wang, T. Gu, W. Yao, Research on the application of the Sleep EEG Net model based on domain adaptation transfer in the detection of driving fatigue, Biomed. Signal. Process. Control 90 (2024), https://doi.org/10.1016/j.bspc.2023.105832.

[118] H. Zeng, et al., Instanceeasytl: an improved transfer-learning method for EEG-based cross-subject fatigue detection, Sensors (Switzerland) 20 (24) (2020) 1–17, https://doi.org/10.3390/s20247251.

[119] K. Guan, Z. Zhang, T. Liu, H. Niu, Cross-task mental workload recognition based on EEG tensor representation and transfer learning, IEEE Trans. Neural Syst. Rehabil. Eng. 31 (2023) 2632–2639, https://doi.org/10.1109/TNSRE.2023.3277867.

[120] F. Wang, W. Zhang, Z. Xu, J. Ping, H. Chu, A deep multi-source adaptation transfer network for cross-subject electroencephalogram emotion recognition, Neural Comput. Appl. 33 (15) (2021) 9061–9073, https://doi.org/10.1007/s00521-020-05670-4.

[121] F. Liu, et al., Emotion recognition from few-channel EEG signals by integrating deep feature aggregation and transfer learning, IEEe Trans. Affect. Comput. (2023), https://doi.org/10.1109/TAFFC.2023.3336531.

[122] M.S. Shahabi, A. Shalbaf, Prediction of treatment outcome in major depressive disorder using ensemble of hybrid transfer learning and long short-term memory based on EEG signal, IEEe Trans. Cogn. Dev. Syst. 15 (3) (Sep. 2023) 1279–1288, https://doi.org/10.1109/TCDS.2022.3207350.

[123] M.S. Shahabi, B. Nobakhsh, A. Shalbaf, R. Rostami, R. Kazemi, Prediction of treatment outcome for repetitive transcranial magnetic stimulation in major depressive disorder using connectivity measures and ensemble of pre-trained deep learning models, Biomed. Signal. Process. Control 85 (2023), https://doi.org/10.1016/j.bspc.2023.104822.

[124] M.S. Shahabi, A. Shalbaf, A. Maghsoudi, Prediction of drug response in major depressive disorder using ensemble of transfer learning with convolutional neural network based on EEG, Biocybern. Biomed. Eng. 41 (3) (2021) 946–959, https://doi.org/10.1016/j.bbe.2021.06.006.

[125] Z. He, et al., Cross-scenario automatic sleep stage classification using transfer learning and single-channel EEG, Biomed. Signal. Process. Control 81 (2023), https://doi.org/10.1016/j.bspc.2022.104501.

[126] S. Li, I. Daly, C. Guan, A. Cichocki, J. Jin, Inter-participant transfer learning with attention based domain adversarial training for P300 detection, Neural Netw. 180 (2024), https://doi.org/10.1016/j.neunet.2024.106655.

[127] F. Wei, X. Xu, X. Li, X. Wu, BDAN-SPD: a brain decoding adversarial network guided by spatiotemporal pattern differences for cross-subject MI-BCI, IEEe Trans. Industr. Inform. (2024), https://doi.org/10.1109/TII.2024.3450010.

[128] S. Asgarian, et al., Multiview contrastive learning for unsupervised domain adaptation in brain-computer interfaces, IEEe Trans. Instrum. Meas. 73 (2024) 1–10, https://doi.org/10.1109/TIM.2024.3366285.

[129] J. Zhu, G. Xu, Q. Hu, B. Wang, T. Zhou, J. Qin, Dual contrastive training and transferability-aware adaptation for multisource privacy-preserving motor imagery classification, IEEe Trans. Instrum. Meas. 73 (2024) 1–13, https://doi.org/10.1109/TIM.2023.3341121.

[130] J. Huang, et al., Cross-subject transfer method based on domain generalization for facilitating calibration of SSVEP-based BCIs, IEEE Trans. Neural Syst. Rehabil. Eng. 31 (2023) 3307–3319, https://doi.org/10.1109/TNSRE.2023.3305202.

[131] A.M. Roy, Adaptive transfer learning-based multiscale feature fused deep convolutional neural network for EEG MI multiclassification in brain–computer interface, Eng. Appl. Artif. Intell. 116 (2022) 105347, https://doi.org/10.1016/j.engappai.2022.105347.
[132] J. Li, Q. She, F. Fang, Y. Chen, Y. Zhang, Multi-source transfer learning via optimal transport feature ranking for EEG classification, Neurocomputing. 596 (Sep. 2024), https://doi.org/10.1016/j.neucom.2024.127944.
[133] F. Zhang, H. Wu, Y. Guo, Semi-supervised multi-source transfer learning for cross-subject EEG motor imagery classification, Med. Biol. Eng. Comput. 62 (6) (2024) 1655–1672, https://doi.org/10.1007/s11517-024-03032-z.
[134] Y. Xie, et al., Cross-dataset transfer learning for motor imagery signal classification via multi-task learning and pre-training, J. Neural Eng. 2023 (10) (2023), https://doi.org/10.1088/1741-2552/acfe9c.
[135] M. Miao, Z. Yang, Z. Sheng, B. Xu, W. Zhang, X. Cheng, Multi-source deep domain adaptation ensemble framework for cross-dataset motor imagery EEG transfer learning, Physiol. Meas. 45 (5) (2024), https://doi.org/10.1088/1361-6579/ad4e95.
[136] H.Y. Yap, Y.H. Choo, Z.I. Mohd Yusoh, W.H. Khoh, An evaluation of transfer learning models in EEG-based authentication, Brain Inform. 10 (1) (2023), https://doi.org/10.1186/s40708-023-00198-4.
[137] W. Qu, et al., Single-channel EEG based insomnia detection with domain adaptation, Comput. Biol. Med. 139 (2021), https://doi.org/10.1016/j.compbiomed.2021.104989.
[138] S. Siuly, S.K. Khare, E. Kabir, M.T. Sadiq, H. Wang, An efficient Parkinson's disease detection framework: leveraging time-frequency representation and AlexNet convolutional neural network, Comput. Biol. Med. 174 (2024), https://doi.org/10.1016/j.compbiomed.2024.108462.
[139] M.N. Islam, N. Sulaiman, F. Al Farid, J. Uddin, S.A. Alyami, M. Rashid, M. A. Moni, Diagnosis of hearing deficiency using EEG based AEP signals: CWT and improved-VGG16 pipeline, PeerJ Computer Science 7 (2021) e638.
[140] Z. Song, B. Deng, J. Wang, G. Yi, An EEG-based systematic explainable detection framework for probing and localizing abnormal patterns in Alzheimer's disease, J. Neural Eng. 19 (3) (2022), https://doi.org/10.1088/1741-2552/ac697d.
[141] H. Bechinia, D. Benmerzoug, N. Khlifa, Approach based lightweight custom convolutional neural network and fine-tuned MobileNet-V2 for ECG arrhythmia signals classification, IEEe Access. 12 (2024) 40827–40841, https://doi.org/10.1109/ACCESS.2024.3378730.
[142] S. Nahak, A. Pathak, G. Saha, Evaluation of handcrafted features and learned representations for the classification of arrhythmia and congestive heart failure in ECG, Biomed. Signal. Process. Control 79 (2023), https://doi.org/10.1016/j.bspc.2022.104230.
[143] Z. He, et al., A novel unsupervised domain adaptation framework based on graph convolutional network and multi-level feature alignment for inter-subject ECG classification, Expert. Syst. Appl. 221 (2023), https://doi.org/10.1016/j.eswa.2023.119711.
[144] M. Du, Y. Yang, L. Zhang, Diagnosis of atrial fibrillation based on unsupervised domain adaptation, Comput. Biol. Med. 164 (2023), https://doi.org/10.1016/j.compbiomed.2023.107275.
[145] Y. Ng, M.T. Liao, T.L. Chen, C.K. Lee, C.Y. Chou, W. Wang, Few-shot transfer learning for personalized atrial fibrillation detection using patient-based siamese network with single-lead ECG records, Artif. Intell. Med. 144 (2023), https://doi.org/10.1016/j.artmed.2023.102644.
[146] N. Gahungu, A. Shariar, D. Playford, C. Judkins, E. Gabbay, Transfer learning artificial intelligence for automated detection of atrial fibrillation in patients undergoing evaluation for suspected obstructive sleep apnoea: a feasibility study, Sleep. Med. 85 (2021) 166–171, https://doi.org/10.1016/j.sleep.2021.07.014.
[147] H. Xu, et al., A dynamic transfer network for cross-database atrial fibrillation detection, Biomed. Signal. Process. Control 90 (2024), https://doi.org/10.1016/j.bspc.2023.105799.
[148] I. Moran, D.T. Altilar, M.K. Ucar, C. Bilgin, M.R. Bozkurt, Deep transfer learning for chronic obstructive pulmonary disease detection utilizing electrocardiogram signals, IEEe Access. 11 (2023) 40629–40644, https://doi.org/10.1109/ACCESS.2023.3269397.
[149] S. Bhaskarpandit, A. Gade, S. Dash, D.K. Dash, R.K. Tripathy, R.B. Pachori, Detection of myocardial infarction from 12-lead ECG trace images using Eigendomain deep representation learning, IEEe Trans. Instrum. Meas. 72 (2023), https://doi.org/10.1109/TIM.2023.3241986.
[150] Y. Gong, L. Wei, S. Yan, F. Zuo, H. Zhang, Y. Li, Transfer learning based deep network for signal restoration and rhythm analysis during cardiopulmonary resuscitation using only the ECG waveform, Inf. Sci. (N. Y.) 626 (2023) 754–772, https://doi.org/10.1016/j.ins.2023.01.055.
[151] S. Ishaque, N. Khan, S. Krishnan, Comprehending the impact of deep learning algorithms on optimizing for recurring impediments associated with stress prediction using ECG data through statistical analysis, Biomed. Signal. Process. Control 74 (2022), https://doi.org/10.1016/j.bspc.2022.103484.
[152] O. Yildirim, M. Talo, B. Ay, U.B. Baloglu, G. Aydin, U.R. Acharya, Automated detection of diabetic subject using pre-trained 2D-CNN models with frequency spectrum images extracted from heart rate signals, Comput. Biol. Med. 113 (2019), https://doi.org/10.1016/j.compbiomed.2019.103387.
[153] P. Sarkar and A. Etemad, "Self-supervised ECG representation learning for emotion recognition," Feb. 2020, doi: 10.1109/TAFFC.2020.3014842.
[154] S. Rahman, S. Pal, J. Yearwood, C. Karmakar, Robustness of deep learning models in electrocardiogram noise detection and classification, Comput. Methods Programs Biomed. 253 (2024), https://doi.org/10.1016/j.cmpb.2024.108249.
[155] K. Prashant, P. Choudhary, T. Agrawal, E. Kaushik, OWAE-Net: COVID-19 detection from ECG images using deep learning and optimized weighted average ensemble technique, Intell. Syst. Appl. 16 (2022), https://doi.org/10.1016/j.iswa.2022.200154.
[156] R. Srivastva, A. Singh, Y.N. Singh, PlexNet: a fast and robust ECG biometric system for human recognition, Inf. Sci. (N. Y.) 558 (2021) 208–228, https://doi.org/10.1016/j.ins.2021.01.001.
[157] Y.H. Byeon, S.B. Pan, K.C. Kwak, Intelligent deep models based on scalograms of electrocardiogram signals for biometrics, Sensors (Switzerland) 19 (4) (2019), https://doi.org/10.3390/s19040935.
[158] K. Suvila, T.J. Niiranen, Interrelations between high blood pressure, organ damage, and cardiovascular disease: no more room for doubt, Hypertension 79 (3) (2022) 516–517, https://doi.org/10.1161/HYPERTENSIONAHA.121.18786.
[159] A. Moussouni, A. Sidi-yakhlef, H. Hamdaoui, A. Aouar, D. Belkhatir, Prevalence and risk factors of prehypertension and hypertension in Algeria, BMC. Public Health 22 (1) (2022) 1571, https://doi.org/10.1186/s12889-022-13942-y.
[160] M. Cejnar, H. Kobler, S.N. Hunyor, Quantitative photoplethysmography: Lambert-Beer law or inverse function incorporating light scatter, J. Biomed. Eng. 15 (2) (1993) 151–154, https://doi.org/10.1016/0141-5425(93)90047-3.
[161] P. Osathitporn, et al., RRWaveNet: a compact end-to-end multiscale residual CNN for robust PPG Respiratory rate estimation, IEEe Internet. Things. J. 10 (18) (2023) 15943–15952, https://doi.org/10.1109/JIOT.2023.3265980.
[162] H.M. Koparır, Ö. Arslan, Cuffless blood pressure estimation from photoplethysmography using deep convolutional neural network and transfer learning, Biomed. Signal. Process. Control 93 (2024), https://doi.org/10.1016/j.bspc.2024.106194.
[163] K. Qin, W. Huang, T. Zhang, Deep generative model with domain adversarial training for predicting arterial blood pressure waveform from photoplethysmogram signal, Biomed. Signal. Process. Control 70 (2021), https://doi.org/10.1016/j.bspc.2021.102972.
[164] C. Ma, et al., STP: self-supervised transfer learning based on transformer for noninvasive blood pressure estimation using photoplethysmography, Expert. Syst. Appl. 249 (2024), https://doi.org/10.1016/j.eswa.2024.123809.
[165] R. Song, S. Zhang, C. Li, Y. Zhang, J. Cheng, X. Chen, Heart rate estimation from facial videos using a spatiotemporal representation with convolutional neural networks, IEEe Trans. Instrum. Meas. 69 (10) (2020) 7411–7421, https://doi.org/10.1109/TIM.2020.2984168.
[166] S. Zanelli, M.A. El Yacoubi, M. Hallab, M. Ammi, Type 2 diabetes detection with light CNN from single raw PPG wave, IEEe Access. 11 (2023) 57652–57665, https://doi.org/10.1109/ACCESS.2023.3274484.
[167] A.K. Abbas, R. Bassam, Phonocardiography Signal Processing, Springer International Publishing, Cham, 2009, https://doi.org/10.1007/978-3-031-01637-0.
[168] S.A. Singh, N.D. Devi, K.N. Singh, K. Thongam, R.D. B, S. Majumder, An ensemble-based transfer learning model for predicting the imbalance heart sound signal using spectrogram images, Multimed. Tools. Appl. 83 (13) (2024) 39923–39942, https://doi.org/10.1007/s11042-023-17186-9.
[169] J. Karhade, S. Dash, S.K. Ghosh, D.K. Dash, R.K. Tripathy, Time-frequency-domain deep learning framework for the automated detection of heart valve disorders using PCG signals, IEEe Trans. Instrum. Meas. 71 (2022), https://doi.org/10.1109/TIM.2022.3163156.
[170] Y. Zheng, X. Guo, Y. Yang, H. Wang, K. Liao, J. Qin, Phonocardiogram transfer learning-based CatBoost model for diastolic dysfunction identification using multiple domain-specific deep feature fusion, Comput. Biol. Med. 156 (2023), https://doi.org/10.1016/j.compbiomed.2023.106707.
[171] C.Y. Yang, P.C. Chen, W.C. Huang, Cross-domain transfer of EEG to EEG or ECG learning for CNN classification models, Sensors 23 (5) (2023), https://doi.org/10.3390/s23052458.
[172] D. Calhas, R. Henriques, Eeg to fmri synthesis for medical decision support: A case study on schizophrenia diagnosis, medRxiv (2023), 2023...08.
[173] J.J. Bird, J. Kobylarz, D.R. Faria, A. Ekart, E.P. Ribeiro, Cross-domain MLP and CNN transfer learning for biological signal processing: EEG and EMG, IEEe Access. 8 (2020) 54789–54801, https://doi.org/10.1109/ACCESS.2020.2979074.
[174] K. Uludağ, A. Roebroeck, General overview on the merits of multimodal neuroimaging data fusion, Neuroimage 102 (2014) 3–10, https://doi.org/10.1016/j.neuroimage.2014.05.018.
[175] M. Ewers et al., "Staging Alzheimer's disease progression with multimodality neuroimaging," Dec. 2011. doi: 10.1016/j.pneurobio.2011.06.004.
[176] C.J. Aine, et al., Multimodal neuroimaging in Schizophrenia: description and dissemination, Neuroinformatics. 15 (4) (2017) 343–364, https://doi.org/10.1007/s12021-017-9338-9.
[177] P.V. J, A.V.S. A, K.P. G, N.K. Karthikeyan, A novel attention-based cross-modal transfer learning framework for predicting cardiovascular disease, Comput. Biol. Med. 170 (2024), https://doi.org/10.1016/j.compbiomed.2024.107977.
[178] M. Abdollahpour, T.Y. Rezaii, A. Farzamnia, I. Saad, Transfer learning convolutional neural network for sleep stage classification using two-stage data fusion framework, IEEe Access. 8 (2020) 180618–180632, https://doi.org/10.1109/ACCESS.2020.3027289.
[179] Y.H. Cheng, M. Lech, R.H. Wilkinson, Distributed Neural network system for multimodal sleep stage detection, IEEe Access. 11 (2023) 29048–29061, https://doi.org/10.1109/ACCESS.2023.3260215.
[180] A. Giovannetti, et al., Deep-MEG: spatiotemporal CNN features and multiband ensemble classification for predicting the early signs of Alzheimer's disease with magnetoencephalography, Neural Comput. Appl. 33 (21) (2021) 14651–14667, https://doi.org/10.1007/s00521-021-06105-4.

[181] Jin, K., Rubio-Solis, A., Naik, R., Onyeogulu, T., Islam, A., Khan, S., ... & Mylonas, G. (2022). Identification of Cognitive Workload during Surgical Tasks with Multimodal Deep Learning. *arXiv preprint arXiv:2209.06208*.

[182] D. Farina and F. Negro, "Accessing the neural drive to muscle and translation to neurorehabilitation technologies," 2012. doi: 10.1109/RBME.2012.2183586.

[183] R. Merletti, P. Parker (Eds.), Electromyography, Wiley, 2004, https://doi.org/10.1002/0471678384.

[184] M.A. Ozdemir, D.H. Kisa, O. Guren, A. Akan, Hand gesture classification using time–frequency images and transfer learning based on CNN, Biomed. Signal. Process. Control 77 (2022) 103787, https://doi.org/10.1016/j.bspc.2022.103787.

[185] H. Nodera, Y. Osaki, H. Yamazaki, A. Mori, Y. Izumi, R. Kaji, Deep learning for waveform identification of resting needle electromyography signals, Clin. Neurophysiol. 130 (5) (2019) 617–623, https://doi.org/10.1016/j.clinph.2019.01.024.

[186] Y. Li, W. Zhang, Q. Zhang, N. Zheng, Transfer learning-based muscle activity decoding scheme by low-frequency sEMG for wearable low-cost application, IEEe Access. 9 (2021) 22804–22815, https://doi.org/10.1109/ACCESS.2021.3056412.

[187] Phoo, C. P., & Hariharan, B. (2020). Self-training for few-shot transfer across extreme task differences. *arXiv preprint arXiv:2010.07734*.

[188] Y. Zou, L. Cheng, A transfer learning model for gesture recognition based on the deep features extracted by CNN, IEEe Trans. Artif. Intell. 2 (5) (2021) 447–458, https://doi.org/10.1109/TAI.2021.3098253.

[189] U. Cote-Allard, et al., A transferable adaptive domain adversarial neural network for virtual reality augmented EMG-based gesture recognition, IEEE Trans. Neural Syst. Rehabil. Eng. 29 (2021) 546–555, https://doi.org/10.1109/TNSRE.2021.3059741.

[190] M. El Shayeb, L.A. Topfer, T. Stafinski, L. Pawluk, D. Menon, Diagnostic accuracy of level 3 portable sleep tests versus level 1 polysomnography for sleep-disordered breathing: a systematic review and meta-analysis, CMAJ. Can. Med. Assoc. J. 186 (1) (2014), https://doi.org/10.1503/cmaj.130952.

[191] K. Gupta, V. Bajaj, I.A. Ansari, U.R. Acharya, Hyp-Net: automated detection of hypertension using deep convolutional neural network and Gabor transform techniques with ballistocardiogram signals, Biocybern. Biomed. Eng. 42 (3) (2022) 784–796, https://doi.org/10.1016/j.bbe.2022.06.001.

[192] D. Ma, M. Izzetoglu, R. Holtzer, X. Jiao, Deep learning based walking tasks classification in older adults using fNIRS, IEEE Trans. Neural Syst. Rehabil. Eng. 31 (2023) 3437–3447, https://doi.org/10.1109/TNSRE.2023.3306365.

[193] K. He, X. Zhang, S. Ren, J. Sun, Deep residual learning for image recognition, in: 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), IEEE, 2016, pp. 770–778, https://doi.org/10.1109/CVPR.2016.90.

[194] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," Sep. 2014, [Online]. Available: http://arxiv.org/abs/1409.1556.

[195] J. Cao, D. Hu, Y. Wang, J. Wang, B. Lei, Epileptic classification with deep-transfer-learning-based feature fusion algorithm, IEEe Trans. Cogn. Dev. Syst. 14 (2) (2022) 684–695, https://doi.org/10.1109/TCDS.2021.3064228.

[196] A. Krizhevsky, I. Sutskever, G.E. Hinton, ImageNet classification with deep convolutional neural networks, Communications of the ACM 60 (6) (2017) 84–90.

[197] O. Faust, M. Salvi, P.D. Barua, S. Chakraborty, F. Molinari, U.R. Acharya, Issues and limitations on the road to fair and inclusive AI solutions for biomedical challenges, Sensors. (Basel) 25 (1) (2025), https://doi.org/10.3390/s25010205.

[198] G.B. Mensah, et al., Assessing the role Ghana's Public Health Act, 2012 (Act 851) can play in oversight of artificial intelligence healthcare systems to prevent medical errors and improve patient safety, Babylonian J. Artif. Intell. 2023 (2023) 24–32, https://doi.org/10.58496/BJAI/2023/006.

[199] L. Alzubaidi, et al., Novel transfer learning approach for medical imaging with limited labeled data, Cancers. (Basel) 13 (7) (2021), https://doi.org/10.3390/cancers13071590.

[200] L.R. Novick, Analogical transfer, problem similarity, and expertise, J. Exp. Psychol. Learn. Mem. Cogn. 14 (3) (1988) 510–520, https://doi.org/10.1037/0278-7393.14.3.510.

[201] L. Ge, J. Gao, H. Ngo, K. Li, A. Zhang, On handling negative transfer and imbalanced distributions in multiple source transfer learning, Stat. Anal. Data Min.: The ASA Data Sci. J. 7 (4) (2014) 254–271, https://doi.org/10.1002/sam.11217.

[202] J.M. Johnson, T.M. Khoshgoftaar, Survey on deep learning with class imbalance, J. Big. Data 6 (1) (2019), https://doi.org/10.1186/s40537-019-0192-5.

[203] D. Hazra, Y.-C. Byun, SynSigGAN: generative adversarial networks for synthetic biomedical signal generation, Biology. (Basel) 9 (12) (2020) 441, https://doi.org/10.3390/biology9120441.

[204] A. Hernandez-Matamoros, H. Fujita, H. Perez-Meana, A novel approach to create synthetic biomedical signals using BiRNN, Inf. Sci. (N. Y.) 541 (2020) 218–241, https://doi.org/10.1016/j.ins.2020.06.019.

[205] Y.-C. Wang, C.-H. Cheng, A multiple combined method for rebalancing medical data with class imbalances, Comput. Biol. Med. 134 (2021) 104527, https://doi.org/10.1016/j.compbiomed.2021.104527.

[206] P.H. Lee, Resampling methods improve the predictive power of modeling in class-imbalanced datasets, Int. J. Environ. Res. Public Health 11 (9) (2014) 9776–9789, https://doi.org/10.3390/ijerph110909776.

[207] S.H. Khan, M. Hayat, M. Bennamoun, F.A. Sohel, R. Togneri, Cost-sensitive learning of deep feature representations from imbalanced data, IEEe Trans. Neural Netw. Learn. Syst. 29 (8) (2018) 3573–3587, https://doi.org/10.1109/TNNLS.2017.2732482.

[208] T. Mao, C. Li, Y. Zhao, R. Song, X. Chen, EEG-based seizure prediction via GhostNet and imbalanced learning, IEEe Sens. Lett. 7 (12) (2023) 1–4, https://doi.org/10.1109/LSENS.2023.3330327.

[209] Z. Liang, Z. Zheng, W. Chen, Z. Pei, J. Wang, J. Chen, A novel deep transfer learning framework integrating general and domain-specific features for EEG-based brain–computer interface, Biomed. Signal. Process. Control 95 (2024), https://doi.org/10.1016/j.bspc.2024.106311.

[210] L. Zhen, P. Hu, X. Peng, R.S.M. Goh, J.T. Zhou, Deep multimodal transfer learning for cross-modal retrieval, IEEe Trans. Neural Netw. Learn. Syst. 33 (2) (2022) 798–810, https://doi.org/10.1109/TNNLS.2020.3029181.

[211] Y.D. Zhang, et al., Advances in multimodal data fusion in neuroimaging: overview, challenges, and novel orientation, Inf. Fusion 64 (2020) 149–187, https://doi.org/10.1016/j.inffus.2020.07.006.

[212] Z. Wang, Z. Dai, B. Póczos, J. Carbonell, Characterizing and avoiding negative transfer, in: *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2019, pp. 11293–11302.

[213] Palanisamy, K., Singhania, D., & Yao, A. (2020). Rethinking CNN models for audio classification. *arXiv preprint arXiv:2007.11154*.

[214] D. Liang, A. Liu, Y. Gao, C. Li, R. Qian, X. Chen, Semi-supervised domain-adaptive seizure prediction via feature alignment and consistency regularization, IEEe Trans. Instrum. Meas. 72 (2023), https://doi.org/10.1109/TIM.2023.3261919.

[215] X. Chai, et al., A fast, efficient domain adaptation technique for cross-domain electroencephalography(EEG)-based emotion recognition, Sensors (Switzerland) 17 (5) (2017), https://doi.org/10.3390/s17051014.

[216] J. Li, S. Qiu, Y.-Y. Shen, C.-L. Liu, H. He, Multisource Transfer Learning for Cross-Subject EEG Emotion Recognition, IEEE Trans Cybern, 2019, pp. 1–13, https://doi.org/10.1109/TCYB.2019.2904052.

[217] E. Dagois, A. Khalaf, E. Sejdic, M. Akcakaya, Transfer learning for a multimodal hybrid EEG-fTCD brain–Computer interface, IEEe Sens. Lett. 3 (1) (2019) 1–4, https://doi.org/10.1109/LSENS.2018.2879466.

[218] J.S. García-Salinas, L. Villaseñor-Pineda, C.A. Reyes-García, A.A. Torres-García, Transfer learning in imagined speech EEG-based BCIs, Biomed. Signal. Process. Control 50 (2019) 151–157, https://doi.org/10.1016/j.bspc.2019.01.006.

[219] M. Du, Y. Yang, L. Zhang, Diagnosis of atrial fibrillation based on unsupervised domain adaptation, Comput. Biol. Med. 164 (2023), https://doi.org/10.1016/j.compbiomed.2023.107275.

[220] X. Jiang, L. Meng, Z. Wang, D. Wu, Deep source semi-supervised transfer learning (DS3TL) for cross-subject EEG classification, IEEE Trans. Biomed. Eng. 71 (4) (2024) 1308–1318, https://doi.org/10.1109/TBME.2023.3333327.

[221] A.G. Roy, S. Siddiqui, S. Pölsterl, N. Navab, and C. Wachinger, "BrainTorrent: a peer-to-peer environment for decentralized federated learning," May 2019, [Online]. Available: http://arxiv.org/abs/1905.06731.

[222] B. Yuan, S. Ge, and W. Xing, "A federated learning framework for healthcare IoT devices," May 2020, [Online]. Available: http://arxiv.org/abs/2005.05083.

[223] N. Gahlan, D. Sethia, Federated learning in emotion recognition systems based on physiological signals for privacy preservation: a review, Multimed. Tools. Appl. (2024), https://doi.org/10.1007/s11042-024-19467-3.

[224] S. Baghersalimi, T. Teijeiro, D. Atienza, A. Aminifar, Personalized real-time federated learning for epileptic seizure detection, IEEe J. Biomed. Health Inform. 26 (2) (2022) 898–909, https://doi.org/10.1109/JBHI.2021.3096127.

[225] A.S. Albahri, et al., A systematic review of trustworthy and explainable artificial intelligence in healthcare: assessment of quality, bias risk, and data fusion, Inf. Fusion 96 (2023) 156–191, https://doi.org/10.1016/j.inffus.2023.03.008.

[226] I. Ahmed, G. Jeon, F. Piccialli, From artificial intelligence to explainable artificial intelligence in industry 4.0: a survey on what, how, and where, IEEe Trans. Industr. Inform. 18 (8) (2022) 5031–5042, https://doi.org/10.1109/TII.2022.3146552.

[227] M.A. Ahmad, A. Teredesai, C. Eckert, Interpretable machine learning in healthcare, in: 2018 IEEE International Conference on Healthcare Informatics (ICHI), IEEE, 2018, https://doi.org/10.1109/ICHI.2018.00095. Jun447–447.

[228] S. Thrun, L. Pratt, Learning to learn: introduction and overview. Learning to Learn, Springer US, Boston, MA, 1998, pp. 3–17, https://doi.org/10.1007/978-1-4615-5529-2_1.

[229] F. Lotte, et al., Computer interfaces: a 10-year update, J. Neural Eng. 15 (3) (2018) 55, https://doi.org/10.1088/1741-2552/aab2f2i.

[230] Q. Wu, et al., Online transfer learning with multiple homogeneous or heterogeneous sources, IEEe Trans. Knowl. Data Eng. 29 (7) (2017) 1494–1507, https://doi.org/10.1109/TKDE.2017.2685597.

[231] Huang, Y., Cheng, P., Tam, R., & Tang, X. (2024). FPT+: A Parameter and Memory Efficient Transfer Learning Method for High-resolution Medical Image Classification. *arXiv preprint arXiv:2408.02426*.

[232] I. Lauriola, A. Lavelli, F. Aiolli, An introduction to deep learning in natural language processing: models, techniques, and tools, Neurocomputing. 470 (2022) 443–456, https://doi.org/10.1016/j.neucom.2021.05.103.

[233] D. Guo, A.M. Rush, and Y. Kim, "Parameter-efficient transfer learning with diff pruning," Dec. 2020, [Online]. Available: http://arxiv.org/abs/2012.07463.

[234] C. Yu, J. Wang, Y. Chen, Z. Wu, Accelerating deep unsupervised domain adaptation with transfer channel pruning, in: *2019 International Joint Conference on Neural Networks (IJCNN)*, IEEE, 2019, July, pp. 1–8.

[235] K.S. Lee, et al., Compressed deep learning to classify arrhythmia in an embedded wearable device, Sensors 22 (5) (2022), https://doi.org/10.3390/s22051776.

[236] X. Liu, A.G. Richardson, Edge deep learning for neural implants: a case study of seizure detection and prediction, J. Neural Eng. 18 (4) (2021), https://doi.org/10.1088/1741-2552/abf473.

[237] Yim, J., Joo, D., Bae, J., & Kim, J. (2017). A gift from knowledge distillation: Fast optimization, network minimization and transfer learning. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 4133-4141).

[238] C. Zhang, Y. Liao, S. Han, M. Zhang, Z. Wang, X. Xie, Multichannel multidomain-based knowledge distillation algorithm for sleep staging with single-channel EEG, IEEE Trans. Circuits Syst. II Express Briefs 69 (11) (2022) 4608–4612, https://doi.org/10.1109/TCSII.2022.3187205.

[239] A.M. Hayajneh, M. Hafeez, S.A.R. Zaidi, D. McLernon, TinyML empowered transfer learning on the edge, IEEE Open J. Commun. Soc. 5 (2024) 1656–1672, https://doi.org/10.1109/OJCOMS.2024.3373177.

[240] E. Tanghatari, M. Kamal, A. Afzali-Kusha, M. Pedram, Distributing DNN training over IoT edge devices based on transfer learning, Neurocomputing. 467 (2022) 56–65, https://doi.org/10.1016/j.neucom.2021.09.045.

[241] H. Zhang, D. Chen, S.B. Ko, New flexible multiple-precision multiply-accumulate unit for deep neural network training and inference, IEEE Trans. Comput. 69 (1) (2020) 26–38, https://doi.org/10.1109/TC.2019.2936192.

[242] R. Krishnan, P. Rajpurkar, E.J. Topol, Self-supervised learning in medicine and healthcare, Nat. Biomed. Eng. 6 (12) (2022) 1346–1352, https://doi.org/10.1038/s41551-022-00914-1.

[243] A. Fedorov, et al., On self-supervised multimodal representation learning: an application to Alzheimer's disease, in: 2021 IEEE 18th International Symposium on Biomedical Imaging (ISBI), IEEE, 2021, pp. 1548–1552, https://doi.org/10.1109/ISBI48211.2021.9434103.

[244] S. Azizi, et al., Robust and data-efficient generalization of self-supervised machine learning for diagnostic imaging, Nat. Biomed. Eng. 7 (6) (2023) 756–779, https://doi.org/10.1038/s41551-023-01049-7.

[245] Seuß, D. (2021). Bridging the gap between explainable AI and uncertainty quantification to enhance trustability. arXiv preprint arXiv:2105.11828.

[246] J. Gawlikowski, et al., A survey of uncertainty in deep neural networks, Artif. Intell. Rev. 56 (2023) 1513–1589, https://doi.org/10.1007/s10462-023-10562-9.

[247] Y. Song, T. Wang, S.K. Mondal, and J.P. Sahoo, "A comprehensive survey of few-shot learning: evolution, applications, challenges, and opportunities," May 2022, [Online]. Available: http://arxiv.org/abs/2205.06743.

[248] M. Ochal, M. Patacchiola, J. Vazquez, A. Storkey, S. Wang, Few-shot learning with class imbalance, IEEe Trans. Artif. Intell. 4 (5) (Oct. 2023) 1348–1358, https://doi.org/10.1109/TAI.2023.3298303.

[249] P. Thuwajit, et al., EEGWaveNet: multiscale CNN-based spatiotemporal feature extraction for EEG seizure detection, IEEe Trans. Industr. Inform. 18 (8) (2022) 5547–5557, https://doi.org/10.1109/TII.2021.3133307.

[250] T. Dissanayake, T. Fernando, S. Denman, S. Sridharan, C. Fookes, Deep learning for patient-independent epileptic seizure prediction using scalp EEG signals, IEEe Sens. J. 21 (7) (2021) 9377–9388, https://doi.org/10.1109/JSEN.2021.3057076.

[251] D. Hu, J. Cao, X. Lai, Y. Wang, S. Wang, Y. Ding, Epileptic State classification by fusing hand-crafted and deep learning EEG features, IEEE Trans. Circuits Syst. II Express Briefs 68 (4) (2021) 1542–1546, https://doi.org/10.1109/TCSII.2020.3031399.

[252] K. Rasheed, J. Qadir, T.J. O'Brien, L. Kuhlmann, A. Razi, A generative model to synthesize EEG data for epileptic seizure prediction, IEEE Trans. Neural Syst. Rehabil. Eng. 29 (2021) 2322–2332, https://doi.org/10.1109/TNSRE.2021.3125023.

[253] H. Daoud, M.A. Bayoumi, Efficient epileptic seizure prediction based on deep learning, IEEe Trans. Biomed. Circuits. Syst. 13 (5) (2019) 804–813, https://doi.org/10.1109/TBCAS.2019.2929053.

[254] S. Goel, R. Agrawal, R.K. Bharti, Automated detection of epileptic EEG signals using recurrence plots-based feature extraction with transfer learning, Soft. comput. (2023), https://doi.org/10.1007/s00500-023-08386-4.

[255] A. Narin, Detection of focal and non-focal epileptic seizure using continuous wavelet transform-based scalogram images and pre-trained Deep Neural networks, IRBM 43 (1) (2022) 22–31, https://doi.org/10.1016/j.irbm.2020.11.002.

[256] P. Peng, L. Xie, K. Zhang, J. Zhang, L. Yang, H. Wei, Domain adaptation for epileptic EEG classification using adversarial learning and riemannian manifold, Biomed. Signal. Process. Control 75 (2022), https://doi.org/10.1016/j.bspc.2022.103555.

[257] C. Yang, Z. Deng, K.S. Choi, S. Wang, Takagi-Sugeno-Kang transfer learning fuzzy logic system for the adaptive recognition of epileptic electroencephalogram signals, IEEE Trans. Fuzzy Syst. 24 (5) (2016) 1079–1094, https://doi.org/10.1109/TFUZZ.2015.2501438.

[258] Y. Jiang, et al., Seizure classification from EEG signals using transfer learning, semi-supervised learning and TSK fuzzy system, IEEE Trans. Neural Syst. Rehabil. Eng. 25 (12) (2017) 2270–2284, https://doi.org/10.1109/TNSRE.2017.2748388.

[259] F. Pisano, et al., Convolutional neural network for seizure detection of nocturnal frontal lobe epilepsy, Complexity. 2020 (2020), https://doi.org/10.1155/2020/4825767.

[260] Z. Shi, Z. Liao, H. Tabata, Enhancing performance of convolutional neural network-based epileptic electroencephalogram diagnosis by asymmetric stochastic resonance, IEEe J. Biomed. Health Inform. 27 (9) (2023) 4228–4239, https://doi.org/10.1109/JBHI.2023.3282251.

[261] Z. Jiang, F.L. Chung, S. Wang, Recognition of multiclass epileptic EEG signals based on knowledge and label space inductive transfer, IEEE Trans. Neural Syst. Rehabil. Eng. 27 (4) (2019) 630–642, https://doi.org/10.1109/TNSRE.2019.2904708.

[262] Y. Jiang, et al., Smart diagnosis: a multiple-source transfer TSK fuzzy system for EEG seizure identification, ACM Trans. Multimed. Comput. Commun. Appl. 16 (2s) (2020), https://doi.org/10.1145/3340240.

[263] A.A. Ein Shoka, M.M. Dessouky, A. El-Sayed, E. El-Din Hemdan, An efficient CNN based epileptic seizures detection framework using encrypted EEG signals for secure telemedicine applications, Alex. Eng. J. 65 (2023) 399–412, https://doi.org/10.1016/j.aej.2022.10.014.

[264] X. Cui, T. Wang, X. Lai, T. Jiang, F. Gao, J. Cao, Cross-subject seizure detection by Joint-probability-discrepancy-based domain adaptation, IEEe Trans. Instrum. Meas. 72 (2023), https://doi.org/10.1109/TIM.2023.3248101.

[265] S. Hu, et al., Exploring the applicability of transfer learning and feature engineering in epilepsy prediction using hybrid transformer model, IEEE Trans. Neural Syst. Rehabil. Eng. 31 (2023) 1321–1332, https://doi.org/10.1109/TNSRE.2023.3244045.

[266] Y. Yao, Z. Cui, Y.Z. Jiang, An automatic epilepsy detection method based on improved inductive transfer learning, Comput. Math. Methods Med. 2020 (2020), https://doi.org/10.1155/2020/5046315.

[267] S. Pattnaik, B.N. Rao, N.K. Rout, S.K. Sabut, Transfer learning based epileptic seizure classification using scalogram images of EEG signals, Multimed. Tools. Appl. (2024), https://doi.org/10.1007/s11042-024-19129-4.

[268] M. Nasseri, et al., Non-invasive wearable seizure detection using long-short-term memory networks with transfer learning, J. Neural Eng. 18 (5) (2021), https://doi.org/10.1088/1741-2552/abef8a.

[269] H. Xu, H. He, W. Xue, Z. Dai, Y. Hao, Transfer learning and clustering analysis of epileptic EEG signals on riemannian manifold, Appl. Soft. Comput. 146 (2023), https://doi.org/10.1016/j.asoc.2023.110656.

[270] A. Li, et al., Multiview transfer representation learning with TSK fuzzy system for EEG epilepsy detection, IEEE Trans. Fuzzy Syst. 32 (1) (2024) 38–52, https://doi.org/10.1109/TFUZZ.2023.3288479.

[271] Z. Zhang, A. Liu, Y. Gao, X. Cui, R. Qian, X. Chen, Distilling invariant representations with domain adversarial learning for cross-subject children seizure prediction, IEEe Trans. Cogn. Dev. Syst. 16 (1) (2024) 202–211, https://doi.org/10.1109/TCDS.2023.3257055.

[272] T. Mao, C. Li, R. Song, G. Xu, X. Chen, Online seizure prediction via fine-tuning and test-time adaptation, IEEe Internet. Things. J. 11 (11) (2024) 20784–20796, https://doi.org/10.1109/JIOT.2024.3373616.

[273] S. Fussner, A. Boyne, A. Han, L.A. Nakhleh, Z. Haneef, Differentiating epileptic and psychogenic non-epileptic seizures using machine learning analysis of EEG plot images, Sensors 24 (9) (2024), https://doi.org/10.3390/s24092823.

[274] L. Wei, J.C. Mchugh, C. Mooney, Transfer learning for the identification of paediatric EEGs with interictal epileptiform abnormalities, IEEe Access. 12 (2024) 86073–86082, https://doi.org/10.1109/ACCESS.2024.3415786.

[275] F. Lopes, M.F. Pinto, A. Dourado, A. Schulze-Bonhage, M. Dümpelmann, C. Teixeira, Addressing data limitations in seizure prediction through transfer learning, Sci. Rep. 14 (1) (2024), https://doi.org/10.1038/s41598-024-64802-1.

[276] S.K. Khare, V. Bajaj, U.R. Acharya, SPWVD-CNN for automated detection of schizophrenia patients using EEG signals, IEEe Trans. Instrum. Meas. 70 (2021), https://doi.org/10.1109/TIM.2021.3070608.

[277] A. Shalbaf, S. Bagherzadeh, A. Maghsoudi, Transfer learning with deep convolutional neural network for automated detection of schizophrenia from EEG signals, Phys. Eng. Sci. Med. 43 (4) (2020) 1229–1239, https://doi.org/10.1007/s13246-020-00925-9.

[278] A.I. Korda, E. Ventouras, P. Asvestas, M. Toumaian, G.K. Matsopoulos, N. Smyrnis, Convolutional neural network propagation on electroencephalographic scalograms for detection of schizophrenia, Clin. Neurophysiol. 139 (2022) 90–105, https://doi.org/10.1016/j.clinph.2022.04.010.

[279] N. Sobahi, B. Ari, H. Cakar, O.F. Alcin, A. Sengur, A new signal to image mapping procedure and convolutional neural networks for efficient schizophrenia detection in EEG recordings, IEEe Sens. J. 22 (8) (2022) 7913–7919, https://doi.org/10.1109/JSEN.2022.3151465.

[280] N.K. Al-Qazzaz, A.A. Aldoori, A.K. Buniya, S.H.B.M. Ali, S.A. Ahmad, Transfer learning and hybrid deep convolutional neural networks models for Autism spectrum disorder classification from EEG signals, IEEe Access. 12 (2024) 64510–64530, https://doi.org/10.1109/ACCESS.2024.3396869.

[281] R.S. Lalawat, V. Bajaj, An automatic framework for detecting Autism spectrum disorder from EEG signals using TFD, IEEe Sens. J. 24 (7) (2024) 10632–10639, https://doi.org/10.1109/JSEN.2024.3362341.

[282] J. Chen, H. Wang, E. He, A transfer learning-based CNN deep learning model for unfavorable driving State recognition, Cognit. Comput. 16 (1) (2024) 121–130, https://doi.org/10.1007/s12559-023-10196-7.

[283] H. Zeng, et al., An eeg-based transfer learning method for cross-subject fatigue mental state prediction, Sensors 21 (7) (2021), https://doi.org/10.3390/s21072369.

[284] C. Ma, M. Zhang, X. Sun, H. Wang, Z. Gao, Dynamic threshold distribution domain adaptation Network: a cross-subject fatigue recognition method based on EEG signals, IEEe Trans. Cogn. Dev. Syst. 16 (1) (2024) 190–201, https://doi.org/10.1109/TCDS.2023.3257428.

[285] Y. Gu, K. Xia, K.W. Lai, Y. Jiang, P. Qian, X. Gu, Transferable Takagi-Sugeno-Kang fuzzy classifier with multi-views for EEG-based driving fatigue recognition in intelligent transportation, IEEE Trans. Intell. Transp. Syst. 24 (12) (2023) 15807–15817, https://doi.org/10.1109/TITS.2022.3220597.

[286] Z. Yin, M. Zhao, W. Zhang, Y. Wang, Y. Wang, J. Zhang, Physiological-signal-based mental workload estimation via transfer dynamical autoencoders in a deep learning framework, Neurocomputing. 347 (2019) 212–229, https://doi.org/10.1016/j.neucom.2019.02.061.

[287] L. Jin, H. Qu, L. Pang, Z. Zhang, Z. Lyu, Identifying stable EEG patterns over time for mental workload recognition using transfer DS-CNN framework, Biomed. Signal. Process. Control 89 (2024), https://doi.org/10.1016/j.bspc.2023.105662.

[288] S. Yang, A. Shan, L. Wang, Y. Li, S. Liu, Cross-subject and cross-experimental classification of mental fatigue based on two-stream self-attention network, Biomed. Signal. Process. Control 88 (2024), https://doi.org/10.1016/j.bspc.2023.105638.
[289] S. Bagherzadeh, K. Maghooli, A. Shalbaf, A. Maghsoudi, Recognition of emotional states using frequency effective connectivity maps through transfer learning approach from electroencephalogram signals, Biomed. Signal. Process. Control 75 (2022), https://doi.org/10.1016/j.bspc.2022.103544.
[290] S. Bagherzadeh, A. Shalbaf, A. Shoeibi, M. Jafari, R.S. Tan, U.R. Acharya, Developing an EEG-based emotion recognition using ensemble deep learning methods and fusion of brain effective connectivity maps, IEEe Access. (2024), https://doi.org/10.1109/ACCESS.2024.3384303.
[291] S. Sartipi, M. Torkamani-Azar, M. Cetin, A hybrid end-to-end spatiotemporal attention neural network with graph-smooth signals for EEG emotion recognition, IEEe Trans. Cogn. Dev. Syst. 16 (2) (2024) 732–743, https://doi.org/10.1109/TCDS.2023.3293321.
[292] S. Liu, Z. Wang, Y. An, B. Li, X. Wang, Y. Zhang, DA-CapsNet: a multi-branch capsule network based on adversarial domain adaption for cross-subject EEG emotion recognition, Knowl. Based. Syst. 283 (2024), https://doi.org/10.1016/j.knosys.2023.111137.
[293] P. Yu, et al., FMLAN: a novel framework for cross-subject and cross-session EEG emotion recognition, Biomed. Signal. Process. Control 100 (2025), https://doi.org/10.1016/j.bspc.2024.106912.
[294] X. Gu, W. Cai, M. Gao, Y. Jiang, X. Ning, P. Qian, Multi-source domain transfer discriminative dictionary learning modeling for electroencephalogram-based emotion recognition, IEEe Trans. Comput. Soc. Syst. 9 (6) (2022) 1604–1612, https://doi.org/10.1109/TCSS.2022.3153660.
[295] W. Li, W. Huan, S. Shao, B. Hou, A. Song, MS-FRAN: a novel multi-source domain adaptation method for EEG-based emotion recognition, IEEe J. Biomed. Health Inform. 27 (11) (2023) 5302–5313, https://doi.org/10.1109/JBHI.2023.3311338.
[296] R. Zhou, et al., PR-PL: a novel prototypical representation based pairwise learning framework for emotion recognition using EEG signals, IEEe Trans. Affect. Comput. (2023), https://doi.org/10.1109/TAFFC.2023.3288118.
[297] J. Li, et al., Cross-subject EEG emotion recognition combined with connectivity features and meta-transfer learning, Comput. Biol. Med. 145 (2022), https://doi.org/10.1016/j.compbiomed.2022.105519.
[298] X. Zhang, et al., Discriminative joint knowledge transfer with online updating mechanism for EEG-based emotion recognition, IEEe Trans. Comput. Soc. Syst. 11 (2) (2024) 2918–2929, https://doi.org/10.1109/TCSS.2023.3314508.
[299] M.B. Er, H. Çiğ, İ.B. Aydilek, A new approach to recognition of human emotions using brain signals and music stimuli, Appl. Acoust. 175 (2021), https://doi.org/10.1016/j.apacoust.2020.107840.
[300] M.S. Shahabi, A. Shalbaf, B. Nobakhsh, R. Rostami, R. Kazemi, Attention-based convolutional recurrent deep neural networks for the prediction of response to repetitive transcranial magnetic stimulation for major depressive disorder, Int. J. Neural Syst. 33 (02) (2023), https://doi.org/10.1142/S0129065723500077.
[301] M.S. Shahabi, A. Shalbaf, R. Rostami, R. Kazemi, A convolutional recurrent neural network with attention for response prediction to repetitive transcranial magnetic stimulation in major depressive disorder, Sci. Rep. 13 (1) (2023), https://doi.org/10.1038/s41598-023-35545-2.
[302] P. Jadhav, G. Rajguru, D. Datta, S. Mukhopadhyay, Automatic sleep stage classification using time–frequency images of CWT and transfer learning using convolution neural network, Biocybern. Biomed. Eng. 40 (1) (2020) 494–504, https://doi.org/10.1016/j.bbe.2020.01.010.
[303] H. ElMoaqet, M. Eid, M. Ryalat, T. Penzel, A deep transfer learning framework for sleep stage classification with single-channel EEG signals, Sensors 22 (22) (2022), https://doi.org/10.3390/s22228826.
[304] H. Wang, H. Guo, K. Zhang, L. Gao, J. Zheng, Automatic sleep staging method of EEG signal based on transfer learning and fusion network, Neurocomputing. 488 (2022) 183–193, https://doi.org/10.1016/j.neucom.2022.02.049.
[305] N.I. Tapia-Rivas, P.A. Estévez, J.A. Cortes-Briones, A robust deep learning detector for sleep spindles and K-complexes: towards population norms, Sci. Rep. 14 (1) (2024), https://doi.org/10.1038/s41598-023-50736-7.
[306] G. Xu, et al., A deep transfer convolutional neural network framework for EEG signal classification, IEEe Access. 7 (2019) 112767–112776, https://doi.org/10.1109/ACCESS.2019.2930958.
[307] T. jian Luo, Dual regularized spatial-temporal features adaptation for multi-source selected cross-subject motor imagery EEG classification, Expert. Syst. Appl. 255 (2024), https://doi.org/10.1016/j.eswa.2024.124673.
[308] J. Lee, J.W. Choi, S. Jo, Selective multi-source domain adaptation network for cross-subject motor imagery discrimination, IEEe Trans. Cogn. Dev. Syst. 16 (3) (2024) 923–934, https://doi.org/10.1109/TCDS.2023.3314351.
[309] A. Nagarajan, N. Robinson, K.K. Ang, K.S.G. Chua, E. Chew, C. Guan, Transferring a deep learning model from healthy subjects to stroke patients in a motor imagery brain–computer interface, J. Neural Eng. 21 (1) (2024) 016007, https://doi.org/10.1088/1741-2552/ad152f.
[310] W. Li, et al., Self-supervised contrastive learning for EEG-based cross-subject motor imagery recognition, J. Neural Eng. 21 (2) (2024), https://doi.org/10.1088/1741-2552/ad3986.
[311] X. Zhang, Z. Miao, C. Menon, Y. Zheng, M. Zhao, D. Ming, Priming cross-session motor imagery classification with a universal deep domain adaptation framework, Neurocomputing. 556 (2023), https://doi.org/10.1016/j.neucom.2023.126659.
[312] D. Li, J. Wang, J. Xu, X. Fang, Y. Ji, Cross-channel specific-mutual feature transfer learning for motor imagery EEG signals decoding, IEEe Trans. Neural Netw. Learn. Syst. (2023), https://doi.org/10.1109/TNNLS.2023.3269512.
[313] R. Zhang, Q. Zong, L. Dou, X. Zhao, Y. Tang, Z. Li, Hybrid deep neural network using transfer learning for EEG motor imagery decoding, Biomed. Signal. Process. Control 63 (2021), https://doi.org/10.1016/j.bspc.2020.102144.
[314] J. Luo, W. Cui, S. Xu, L. Wang, H. Chen, Y. Li, A cross-scale transformer and triple-view attention based domain-rectified transfer learning for EEG classification in RSVP tasks, IEEE Trans. Neural Syst. Rehabil. Eng. 32 (2024) 672–683, https://doi.org/10.1109/TNSRE.2024.3359191.
[315] P. Chen, Z. Gao, M. Yin, J. Wu, K. Ma, C. Grebogi, Multiattention adaptation network for motor imagery recognition, IEEe Trans. Syst. Man. Cybern. Syst. 52 (8) (2022) 5127–5139, https://doi.org/10.1109/TSMC.2021.3114145.
[316] M. Zheng, Y. Lin, A deep transfer learning network with two classifiers based on sample selection for motor imagery brain-computer interface, Biomed. Signal. Process. Control 89 (2024), https://doi.org/10.1016/j.bspc.2023.105786.
[317] E. Jeon, W. Ko, J.S. Yoon, and H.-I. Suk, "Mutual information-driven subject-invariant and class-relevant deep representation learning in BCI," Oct. 2019, [Online]. Available: http://arxiv.org/abs/1910.07747.
[318] Z. Khademi, F. Ebrahimi, H.M. Kordy, A transfer learning-based CNN and LSTM hybrid deep learning model to classify motor imagery EEG signals, Comput. Biol. Med. 143 (2022), https://doi.org/10.1016/j.compbiomed.2022.105288.
[319] E. Santamaria-Vazquez, V. Martinez-Cagigal, F. Vaquerizo-Villar, R. Hornero, EEG-inception: a novel deep convolutional neural network for assistive ERP-based brain-computer interfaces, IEEE Trans. Neural Syst. Rehabil. Eng. 28 (12) (2020) 2773–2782, https://doi.org/10.1109/TNSRE.2020.3048106.
[320] P.R.A.S. Bassi, W. Rampazzo, R. Attux, Transfer learning and SpecAugment applied to SSVEP based BCI classification, Biomed. Signal. Process. Control 67 (2021), https://doi.org/10.1016/j.bspc.2021.102542.
[321] S. Kundu, S. Ari, MsCNN: a deep learning framework for P300-based brain–computer interface speller, IEEe Trans. Med. Robot. Bionics. 2 (1) (2020) 86–93, https://doi.org/10.1109/TMRB.2019.2959559.
[322] F. Mattioli, C. Porcaro, G. Baldassarre, A 1D CNN for high accuracy classification and transfer learning in motor imagery EEG-based brain-computer interface, J. Neural Eng. 18 (6) (2021), https://doi.org/10.1088/1741-2552/ac4430.
[323] X. Zhao, et al., Deep CNN model based on serial-parallel structure optimization for four-class motor imagery EEG classification, Biomed. Signal. Process. Control 72 (2022), https://doi.org/10.1016/j.bspc.2021.103338.
[324] C. Chu, L. Zhu, A. Huang, P. Xu, N. Ying, J. Zhang, Transfer learning with data alignment and optimal transport for EEG based motor imagery classification, J. Neural Eng. 21 (1) (2024), https://doi.org/10.1088/1741-2552/ad1f7a.
[325] Q. She, Y. Cai, S. Du, Y. Chen, Multi-source manifold feature transfer learning with domain selection for brain-computer interfaces, Neurocomputing. 514 (2022) 313–327, https://doi.org/10.1016/j.neucom.2022.09.124.
[326] D.Y. Lee, J.H. Jeong, B.H. Lee, S.W. Lee, Motor imagery classification using inter-task transfer learning via a channel-wise variational autoencoder-based convolutional neural network, IEEE Trans. Neural Syst. Rehabil. Eng. 30 (2022) 226–237, https://doi.org/10.1109/TNSRE.2022.3143836.
[327] L. Zhu, et al., Multi-source fusion domain adaptation using resting-State knowledge for motor imagery classification tasks, IEEe Sens. J. 21 (19) (2021) 21772–21781, https://doi.org/10.1109/JSEN.2021.3101684.
[328] A.G. Lazcano-Herrera, R.Q. Fuentes-Aguilar, A. Ramirez-Morales, M. Alfaro-Ponce, BiLSTM and SqueezeNet with transfer learning for EEG motor imagery classification: validation with own dataset, IEEe Access. 11 (2023) 136422–136436, https://doi.org/10.1109/ACCESS.2023.3328254.
[329] S. Asgarian, et al., Multiview contrastive learning for unsupervised domain adaptation in brain-computer interfaces, IEEe Trans. Instrum. Meas. 73 (2024) 1–10, https://doi.org/10.1109/TIM.2024.3366285.
[330] M. Dai, S. Wang, D. Zheng, R. Na, S. Zhang, Domain transfer multiple kernel boosting for classification of EEG motor imagery signals, IEEe Access. 7 (2019) 49951–49960, https://doi.org/10.1109/ACCESS.2019.2908851.
[331] G. Liang, D. Cao, J. Wang, Z. Zhang, Y. Wu, EISATC-fusion: inception self-attention temporal convolutional network fusion for motor imagery EEG decoding, IEEE Trans. Neural Syst. Rehabil. Eng. 32 (2024) 1535–1545, https://doi.org/10.1109/TNSRE.2024.3382226.
[332] P. Chen, H. Wang, X. Sun, H. Li, C. Grebogi, Z. Gao, Transfer learning with optimal transportation and frequency mixup for EEG-based motor imagery recognition, IEEE Trans. Neural Syst. Rehabil. Eng. 30 (2022) 2866–2875, https://doi.org/10.1109/TNSRE.2022.3211881.
[333] H.W. Ng, C. Guan, Subject-independent meta-learning framework towards optimal training of EEG-based classifiers, Neural Netw. 172 (2024), https://doi.org/10.1016/j.neunet.2024.106108.
[334] X. Wang, R. Yang, M. Huang, An unsupervised deep-transfer-learning-based motor imagery EEG classification scheme for brain–Computer interface, Sensors 22 (6) (2022), https://doi.org/10.3390/s22062241.
[335] M.T. Sadiq, M.Z. Aziz, A. Almogren, A. Yousaf, S. Siuly, A.U. Rehman, Exploiting pretrained CNN models for the development of an EEG-based robust BCI framework, Comput. Biol. Med. 143 (2022), https://doi.org/10.1016/j.compbiomed.2022.105242.
[336] W. Huang, Z. Guan, K. Li, Y. Zhou, Y. Li, An affective brain-computer interface based on a transfer learning method, IEEe Trans. Affect. Comput. (2023), https://doi.org/10.1109/TAFFC.2023.3305982.
[337] D. Wu, "Online and offline domain adaptation for reducing BCI calibration effort," Feb. 2017, doi: 10.1109/THMS.2016.2608931.

[338] Y. Gao, M. Li, Y. Peng, F. Fang, Y. Zhang, Double stage transfer learning for brain-computer interfaces, IEEE Trans. Neural Syst. Rehabil. Eng. 31 (2023) 1128–1136, https://doi.org/10.1109/TNSRE.2023.3241301.
[339] M. Turkoglu, O.F. Alcin, M. Aslan, A. Al-Zebari, A. Sengur, Deep rhythm and long short term memory-based drowsiness detection, Biomed. Signal. Process. Control 65 (2021), https://doi.org/10.1016/j.bspc.2020.102364.
[340] M. Alhussein, G. Muhammad, M.S. Hossain, EEG pathology detection based on deep learning, IEEe Access. 7 (2019) 27781–27788, https://doi.org/10.1109/ACCESS.2019.2901672.
[341] N. Khasawneh, M. Fraiwan, L. Fraiwan, Detection of K-complexes in EEG signals using deep transfer learning and YOLOv3, Cluster. Comput. 26 (6) (2023) 3985–3995, https://doi.org/10.1007/s10586-022-03802-0.
[342] N. Khasawneh, M. Fraiwan, L. Fraiwan, Detection of K-complexes in EEG waveform images using faster R-CNN and deep transfer learning, BMC. Med. Inform. Decis. Mak. 22 (1) (2022), https://doi.org/10.1186/s12911-022-02042-x.
[343] S.K. Prabhakar, S.W. Lee, Improved sparse representation based robust hybrid feature extraction models with transfer and deep learning for EEG classification, Expert. Syst. Appl. 198 (2022), https://doi.org/10.1016/j.eswa.2022.116783.
[344] N. Ilakiyaselvan, A.N. Khan, A. Shahina, Reconstructed phase space portraits for detecting brain diseases using deep learning, Biomed. Signal. Process. Control 71 (2022), https://doi.org/10.1016/j.bspc.2021.103278.
[345] X. Li, X. Tang, J. Liu, Q. Liu, H. Wang, K. Liu, Parameter transfer and riemannian space coordinate alignment for EEG intention recognition, Biomed. Signal. Process. Control 92 (2024), https://doi.org/10.1016/j.bspc.2024.106044.
[346] S. Singh, H. Jadli, R. Padma Priya, V.B. Surya Prasath, KDTL: knowledge-distilled transfer learning framework for diagnosing mental disorders using EEG spectrograms, Neural Comput. Appl. (2024), https://doi.org/10.1007/s00521-024-10207-0.
[347] A. Farahat, C. Reichert, C.M. Sweeney-Reed, H. Hinrichs, Convolutional neural networks for decoding of covert attention focus and saliency maps for EEG feature visualization, J. Neural Eng. 16 (6) (2019), https://doi.org/10.1088/1741-2552/ab3bb4.
[348] B. Muller, R. Lengellé, Cross-gram matrices and their use in transfer learning: application to automatic REM detection using heart rate, Comput. Methods Programs Biomed. 208 (2021), https://doi.org/10.1016/j.cmpb.2021.106280.
[349] R. Fu, Y. Wang, C. Jia, Data augmentation for cross-subject EEG features using Siamese neural network, Biomed. Signal. Process. Control 75 (2022), https://doi.org/10.1016/j.bspc.2022.103614.
[350] N. Bahador, J. Kortelainen, Deep learning-based classification of multichannel bio-signals using directedness transfer learning, Biomed. Signal. Process. Control 72 (2022), https://doi.org/10.1016/j.bspc.2021.103300.
[351] M. Naz, J.H. Shah, M.A. Khan, M. Sharif, M. Raza, R. Damaševičius, From ECG signals to images: a transformation based approach for deep learning, PeerJ. Comput. Sci. 7 (2021) 1–18, https://doi.org/10.7717/PEERJ-CS.386.
[352] C. Li, et al., DeepECG: image-based electrocardiogram interpretation with deep convolutional neural networks, Biomed. Signal. Process. Control 69 (2021), https://doi.org/10.1016/j.bspc.2021.102824.
[353] H. Zhang, et al., MaeFE: masked autoencoders Family of electrocardiogram for self-supervised pretraining and Transfer learning, IEEe Trans. Instrum. Meas. 72 (2023), https://doi.org/10.1109/TIM.2022.3228267.
[354] D.H. Kim, G. Lee, S.H. Kim, An ECG stitching scheme for driver arrhythmia classification based on deep learning, Sensors 23 (6) (2023), https://doi.org/10.3390/s23063257.
[355] H. Ullah, et al., An end-to-end cardiac arrhythmia recognition method with an effective DenseNet model on imbalanced datasets using ECG signal, Comput. Intell. Neurosci. 2022 (2022), https://doi.org/10.1155/2022/9475162.
[356] A.U. Rahman, et al., ECG classification for detecting ECG arrhythmia empowered with deep learning approaches, Comput. Intell. Neurosci. 2022 (2022), https://doi.org/10.1155/2022/6852845.
[357] Y. Cheng, W. Zhu, D. Li, L. Wang, Multi-label classification of arrhythmia using dynamic graph convolutional network based on encoder-decoder framework, Biomed. Signal. Process. Control 95 (2024), https://doi.org/10.1016/j.bspc.2024.106348.
[358] X. Zhang, M. Jiang, K. Polat, A. Alhudhaif, J. Hemanth, W. Wu, Detection of atrial fibrillation from variable-duration ECG signal based on time-adaptive densely network and feature enhancement strategy, IEEe J. Biomed. Health Inform. 27 (2) (2023) 944–955, https://doi.org/10.1109/JBHI.2022.3221464.
[359] X.C. Cao, B. Yao, B.Q. Chen, Atrial fibrillation detection using an improved multi-scale decomposition enhanced residual convolutional neural network, IEEe Access. 7 (2019) 89152–89161, https://doi.org/10.1109/ACCESS.2019.2926749.
[360] Y. Jin, et al., A novel domain adaptive residual network for automatic atrial fibrillation detection, Knowl. Based. Syst. 203 (2020), https://doi.org/10.1016/j.knosys.2020.106122.
[361] M. Deng, et al., An intelligent computer-aided diagnosis method for paroxysmal atrial fibrillation patients with nondiagnostic ECG signals, Biomed. Signal. Process. Control 88 (2024), https://doi.org/10.1016/j.bspc.2023.105683.
[362] S. Liu, J. He, A. Wang, C. Yang, Adaptive atrial fibrillation detection focused on atrial activity analysis, Biomed. Signal. Process. Control 88 (2024), https://doi.org/10.1016/j.bspc.2023.105677.
[363] K. Weimann, T.O.F. Conrad, Transfer learning for ECG classification, Sci. Rep. 11 (1) (2021), https://doi.org/10.1038/s41598-021-84374-8.
[364] F.B. Azam, M.I. Ansari, S.I.S.K. Nuhash, I. McLane, T. Hasan, Cardiac anomaly detection considering an additive noise and convolutional distortion model of heart sound recordings, Artif. Intell. Med. 133 (2022), https://doi.org/10.1016/j.artmed.2022.102417.
[365] R.R. Lopes, et al., Improving electrocardiogram-based detection of rare genetic heart disease using transfer learning: an application to phospholamban p. Arg14del mutation carriers, Comput. Biol. Med. 131 (2021), https://doi.org/10.1016/j.compbiomed.2021.104262.
[366] X. Sun, P. Liu, Z. He, Y. Han, B. Su, Automatic classification of electrocardiogram signals based on transfer learning and continuous wavelet transform, Ecol. Inform. 69 (2022), https://doi.org/10.1016/j.ecoinf.2022.101628.
[367] A. Alghamdi, et al., Detection of myocardial infarction based on novel deep transfer learning methods for urban healthcare in smart cities, Multimed. Tools. Appl. 83 (5) (2024) 14913–14934, https://doi.org/10.1007/s11042-020-08769-x.
[368] M. Hammad, Y. Liu, K. Wang, Multimodal biometric authentication systems using convolution neural network based on different level fusion of ECG and fingerprint, IEEe Access. 7 (2019) 25527–25542, https://doi.org/10.1109/ACCESS.2018.2886573.
[369] A.S. Sakr, P. Pławiak, R. Tadeusiewicz, M. Hammad, Cancelable ECG biometric based on combination of deep transfer learning with DNA and amino acid approaches for human authentication, Inf. Sci. (N. Y.) 585 (2022) 127–143, https://doi.org/10.1016/j.ins.2021.11.066.
[370] M. Amin, et al., ECG-based driver's stress detection using deep transfer learning and fuzzy logic approaches, IEEe Access. 10 (2022) 29788–29809, https://doi.org/10.1109/ACCESS.2022.3158658.
[371] C. Ma, et al., SMART-BP: sEM-ResNet and auto-regressor based on a two-stage framework for noninvasive blood pressure measurement, IEEe Trans. Instrum. Meas. 73 (2024) 1–18, https://doi.org/10.1109/TIM.2023.3342856.
[372] C. Ma, et al., KD-informer: a cuff-less continuous blood pressure waveform estimation approach based on single photoplethysmography, IEEe J. Biomed. Health Inform. 27 (5) (2023) 2219–2230, https://doi.org/10.1109/JBHI.2022.3181328.
[373] Y. Zhang, X. Ren, X. Liang, X. Ye, C. Zhou, A refined blood pressure estimation model based on single channel photoplethysmography, IEEe J. Biomed. Health Inform. 26 (12) (2022) 5907–5917, https://doi.org/10.1109/JBHI.2022.3206477.
[374] F. Schrumpf, P. Frenzel, C. Aust, G. Osterhoff, M. Fuchs, Assessment of non-invasive blood pressure prediction from ppg and rppg signals using deep learning, Sensors 21 (18) (2021), https://doi.org/10.3390/s21186022.
[375] W. Wang, P. Mohseni, K.L. Kilgore, L. Najafizadeh, Cuff-less blood pressure estimation from photoplethysmography via visibility graph and transfer learning, IEEe J. Biomed. Health Inform. 26 (5) (2022) 2075–2085, https://doi.org/10.1109/JBHI.2021.3128383.
[376] J. Leitner, P.H. Chiang, S. Dey, Personalized blood pressure estimation using photoplethysmography: a transfer learning approach, IEEe J. Biomed. Health Inform. 26 (1) (2022) 218–228, https://doi.org/10.1109/JBHI.2021.3085526.
[377] J. Torres-Soto, E.A. Ashley, Multi-task deep learning for cardiac rhythm detection in wearable devices, NPJ. Digit. Med. 3 (1) (2020), https://doi.org/10.1038/s41746-020-00320-4.
[378] A. Maity, A. Pathak, G. Saha, Transfer learning based heart valve disease classification from phonocardiogram signal, Biomed. Signal. Process. Control 85 (2023), https://doi.org/10.1016/j.bspc.2023.104805.
[379] S. Mandala, et al., Enhanced myocardial infarction identification in phonocardiogram signals using segmented feature extraction and transfer learning-based classification, IEEe Access. 11 (2023) 136654–136665, https://doi.org/10.1109/ACCESS.2023.3338853.
[380] S. Ismail, B. Ismail, I. Siddiqi, U. Akram, PCG classification through spectrogram using transfer learning, Biomed. Signal. Process. Control 79 (2022), https://doi.org/10.1016/j.bspc.2022.104075.
[381] A. Pathak, K. Mandana, G. Saha, Ensembled transfer learning and multiple kernel learning for phonocardiogram based atherosclerotic coronary artery disease detection, IEEe J. Biomed. Health Inform. 26 (6) (2022) 2804–2813, https://doi.org/10.1109/JBHI.2022.3140277.
[382] S. Ismail, B. Ismail, PCG signal classification using a hybrid multi round transfer learning classifier, Biocybern. Biomed. Eng. 43 (1) (2023) 313–334, https://doi.org/10.1016/j.bbe.2023.01.004.
[383] P. Dhar, S. Dutta, V. Mukherjee, Cross-wavelet assisted convolution neural network (AlexNet) approach for phonocardiogram signals classification, Biomed. Signal. Process. Control 63 (2021), https://doi.org/10.1016/j.bspc.2020.102142.
[384] O.R.A. Almanifi, A.F. Ab Nasir, M.A. Mohd Razman, R.M. Musa, A.P.P. Abdul Majeed, Heartbeat murmurs detection in phonocardiogram recordings via transfer learning, Alex. Eng. J. 61 (12) (2022) 10995–11002, https://doi.org/10.1016/j.aej.2022.04.031.
[385] A. Harimi, et al., Classification of heart sounds using chaogram transform and deep convolutional neural network transfer learning, Sensors 22 (24) (2022), https://doi.org/10.3390/s22249569.
[386] M. Boulares, T. Alafif, A. Barnawi, Transfer learning benchmark for cardiovascular disease recognition, IEEe Access. 8 (2020) 109475–109491, https://doi.org/10.1109/ACCESS.2020.3002151.
[387] J. Zang, C. Lian, B. Xu, Z. Zhang, Y. Su, C. Xue, AmtNet: attentional multi-scale temporal network for phonocardiogram signal classification, Biomed. Signal. Process. Control 85 (2023), https://doi.org/10.1016/j.bspc.2023.104934.
[388] J. Ramesh, Z. Solatidehkordi, R. Aburukba, A. Sagahyroon, Atrial fibrillation classification with smart wearables using short-term heart rate variability and deep convolutional neural networks, Sensors 21 (21) (2021), https://doi.org/10.3390/s21217233.

[389] J. Zhu, et al., An Improved Convnext with Multimodal Transformer for Physiological Signal Classification, IEEE Access, 2024, https://doi.org/10.1109/ACCESS.2024.3355273.
[390] J. Pan, Y. Yu, J. Wu, X. Zhou, Y. He, Y. Li, Deep neural networks for automatic sleep stage classification and consciousness assessment in patients with disorder of consciousness, IEEe Trans. Cogn. Dev. Syst. (2024), https://doi.org/10.1109/TCDS.2024.3382109.
[391] Y. Lv, J. Xue, F. Duan, Z. Sun, J. Li, An exploratory study of transfer learning frameworks in the context of few available shots of neurophysiological signals, Comput. Electr. Eng. 101 (2022), https://doi.org/10.1016/j.compeleceng.2022.108091.
[392] Siddharth, T. -P. Jung, T. J. Sejnowski, "Utilizing Deep Learning Towards Multi-Modal Bio-Sensing and Vision-Based Affective Computing," in IEEE Transactions on Affective Computing, vol. 13, no. 1, pp. 96-107, 1 Jan.-March 2022, 10.1109/TAFFC.2019.2916015.
[393] E.S. Salama, R.A. El-Khoribi, M.E. Shoman, M.A. Wahby Shalaby, A 3D-convolutional neural network framework with ensemble learning techniques for multi-modal emotion recognition, Egypt. Inform. J. 22 (2) (2021) 167–176, https://doi.org/10.1016/j.eij.2020.07.005.
[394] Y. Huang, J. Yang, S. Liu, J. Pan, Combining facial expressions and electroencephalography to enhance emotion recognition, Future Internet. 11 (5) (2019), https://doi.org/10.3390/fi11050105.
[395] G. Li, et al., Multi-view fusion network-based gesture recognition using sEMG data, IEEe J. Biomed. Health Inform. (2023), https://doi.org/10.1109/JBHI.2023.3287979.
[396] H. Shi, X. Jiang, C. Dai, W. Chen, EMG-based multi-user hand gesture classification via unsupervised transfer learning using unknown calibration gestures, IEEE Trans. Neural Syst. Rehabil. Eng. 32 (2024) 1119–1131, https://doi.org/10.1109/TNSRE.2024.3372002.
[397] K. Khalil, U. Asgher, Y. Ayaz, Novel fNIRS study on homogeneous symmetric feature-based transfer learning for brain–computer interface, Sci. Rep. 12 (1) (2022), https://doi.org/10.1038/s41598-022-06805-4.