Physics-Informed Machine Learning Model Generalization in AIoT: Opportunites and Challenges

Wenjie Luo
School of Computer Science and Engineering
Nanyang Technological University
Singapore

Rui Tan
School of Computer Science and Engineering
Nanyang Technological University
Singapore

Abstract

Recent advances in machine learning inspire the development of deep neural network-based smart sensing applications for the Artificial Intelligence of Things (AIoT). However, due to the nature of the AIoT sensing data, the machine learning models are in general subject to poor generalizability due to the *scarcity of labeled training data* and *run-time domain shifts*. The existing solutions rely on data-driven approaches and do not consider the physical laws that govern data generation or domain shifts. This paper discusses the potential of utilizing the known physical laws to improve the machine learning model generalizability for AIoT applications. Through three case studies, we demonstrate that physics-informed machine learning can (1) effectively assist the generalization of deep neural networks and (2) achieve better performance compared with conventional approaches. Our objective is to encourage more exploration into combining physical principles and machine learning algorithms in physics-rich AIoT.

CCS Concepts

• **Computing methodologies** → **Neural networks**; *Neural networks*; **Spatial and physical reasoning**.

Keywords

Physics-nformed machine learning, Artificial intelligence of things

ACM Reference Format:
Wenjie Luo and Rui Tan. 2022. Physics-Informed Machine Learning Model Generalization in AIoT: Opportunites and Challenges. In *The 3rd Workshop on Data-Driven and Intelligent Cyber-Physical Systems for Smart Cities (DI-CPS), May 9, 2023, San Antonio, Texas USA.* ACM, New York, NY, USA, 6 pages. https://doi.org/xx.xxxx/xxxx.xxxx

1 Introduction

The Internet of Things (IoT) is a global network consisting of trillions of edge sensors. Artificial intelligence (AI) converts vast amounts of IoT sensing data into actionable insights and enhances the smart sensing capabilities of the IoT. The integration of AI, particularly deep learning, with IoT has given rise to the paradigm of Artificial Intelligence of Things (AIoT). Nowadays, AIoT has emerged as an important infrastructure for smart cities, smart transportation, and other smart systems. However, two main challenges arise in developing generalized machine learning models for AIoT sensing, which are

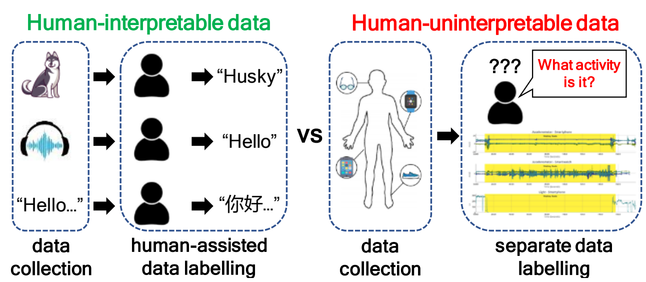


Fig. 1: Human-interpretable media data vs. human-uninterpretable IoT sensing data.

- **Scarcity of labeled training data.** The superior performance of the deep neural network (DNN) models relies on the availability of large, labeled data to uncover useful feature representations. The widespread use of DNNs in computer vision (CV), natural language processing (NLP), and voice sensing can be attributed to the massively available labeled training datasets, such as ImageNet [5] and LibriSpeech corpus [12], etc. Despite the abundance of IoT sensing data, constructing labeled datasets for DNN model training remains a challenging task due to the uninterpretable nature of most sensor data. Fig. 1 uses a simple example to illustrate this challenge. On the left part, it depicts the conventional approach for creating labeled training datasets in the field of CV and NLP. The media data, such as images, audio recordings, and texts, are interpretable by humans. Human assistants can be involved in annotating the data after the data collection is completed. Multiple platforms, such as Amazon Mechanical Turk [1] are established to facilitate this process. On the contrary, as shown in the right part of Fig. 1, the human-uninterpretable property of IoT sensing data makes it difficult for human assistants to understand the meaning behind the data if the labeling process is separated from the data collection. Thus, IoT data labeling must be performed in tandem with data collection, which incurs undesirable overhead, however. The inherent inseparability of the collection and labeling of IoT sensing data results in the scarcity of extensive and labeled training data.
- **Run-time domain shifts** are common in AIoT sensing applications. The domain shifts result in distribution deviations between the collected sensor data at the inference stage and the standard training data. A pre-trained DNN model will

DI-CPS '23, May 9, 2023, San Antonio, Texas USA

ACM ISBN xxx-x-xxxx-xxxx-x/xx/xx.
https://doi.org/xx.xxxx/xxxx.xxxx

have degraded performance if evaluated on collected sensor data due to run-time domain shifts. The cause of these domain shifts can stem from differences in the sensors' hardware or the changing ambient environment where the DNN model is deployed. As a result, the run-time domain shifts need to be carefully addressed while developing DNN models for AIoT sensing applications. In general, the adaptation of a DNN model requires labeled/unlabeled data from the target domain. Thus, data centralization approaches, such as constructing datasets for training generalized DNN models are infeasible in the field of AIoT sensing. Run-time domain shifts exacerbate the scarcity of labeled training data for AIoT sensing applications.

Existing solutions for addressing the above challenges have their limitations. Self-supervised learning, due to its ability in extracting feature representations from large amounts of unlabeled sensing data, has been exploited to address the scarcity of labeled training data in AIoT sensing. Self-attention [17] and contrastive learning [3] are two prevalent techniques. They are used to pre-train DNN models on the abundant unlabeled training data. The pre-trained model is then fine-tuned to a task-specific application using a small amount of labeled training data. Typically, self-supervised learning can reduce the requirement of the labeled training dataset by two orders of magnitude [3, 21]. Despite its capability in feature discovery from unlabeled data, self-supervised learning has its limitations. For example, the quality of learned features can be degraded by noisy or biased data samples. The model trained by self-supervised learning may suffer from overfitting and there is a lack of guidance to evaluate the quality of the feature since labeling information is lacking during the model training. Transfer learning is a prevalent technique to address the run-time domain shift in AIoT sensing. Ideally, transfer learning adapts a pre-trained DNN model from the *source domain* to the *target domain* with limited labeled/unlabeled target-domain data. However, the existing solutions for addressing run-time domain shifts still require a considerable amount of target-domain data [6, 9, 11].

This paper discusses the potential of physics-informed machine learning model generalization in AIoT sensing applications. Specifically, we target the *data scarcity* and *run-time domain shift* challenges and discuss the opportunities and benefits of integrating prior knowledge in the machine learning processes. Through illustrating case studies, we answer the following two questions:

(1) What are the opportunities of the physics-informed machine learning model generalization? In physics-rich IoT sensing, the data generation or domain shifts are governed by certain physical laws. For instance, the distribution of indoor temperature is governed by the fluid dynamics model and the microphone data received is determined by its frequency response curve. A comprehension of these physical laws can aid in the creation of more efficient and generalized machine learning models by incorporating physical laws as additional constraints during model training. The conventional approaches for physical laws estimation involve constructing mathematical models to match the observational data generated by these laws. However, the process of mathematical modeling is often hindered by the need for intricate formulations and costly computations [13]. Recent studies [7, 13] propose physics-informed machine learning (PIML) to leverage the prior knowledge obtained from the observational data to enhance the performance of machine learning models. PIML-based DNN models outperform conventional models and require fewer training data in modeling multi-physics and multi-scale systems.

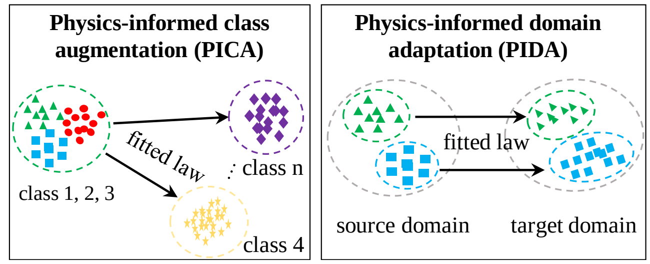


Fig. 2: Left: Physics-informed class augmentation. Right: physics-informed domain adaptation.

In this paper, we examine the potential of using physical laws to enhance the generalization capabilities of machine learning models in AIoT sensing applications. We propose *physics-informed class augmentation (PICA)* and *physics-informed data augmentation (PIDA)* to address the scarcity of labeled training data and run-time domain shifts, respectively. Fig. 2 shows the general ideas of the two approaches. The left part of Fig. 2 depicts PICA, which considers the data scarcity challenge where certain classes and the respective data points are unavailable in the classification tasks. A pre-trained DNN model cannot provide plausible predictions on the data samples out of training classes. To overcome this challenge, PICA takes advantage of the fact that data generation in certain temporal and spatial scales follows a known physical law. This enables us to use the physical law to augment data for the unavailable classes and train a new DNN model that can cover all classes.

The right part of Fig. 2 illustrates PIDA. PIDA considers the AIoT sensing tasks with run-time domain shifts governed by certain physical laws. Typically, obtaining labeled training data for a target AIoT sensing application for DNN model adaptation incurs significant overhead. PIDA addresses the data scarcity in the target domain. Specifically, we identify the physical laws governing the domain shifts and use a small amount of source-domain and target-domain data to fit the physical law. The fitted law is then used to transfer the training dataset from the source domain to the target domain for DNN model adaptation.

(2) What are the benefits of the physics-informed machine learning model generalization? The benefits of the proposed physics-informed machine learning model generalization in AIoT sensing applications are three-fold. First, DNN models have improved generalization capabilities when tested with unfamiliar data samples. PICA uses the fitted physical law to generate synthetic data for the unavailable classes. The trained DNN model can produce accurate predictions when evaluated with data collected from these unavailable classes during training. In §3.1, we use the fitted floor slowness model to generate an augmented dataset for DNN model training. The new location recognition model has reduced recognition errors on data samples generated from unknown locations. Second, incorporating physical principles into machine learning algorithms reduces the need for large amounts of labeled data for

AIoT sensing applications. Both PICA and PIDA use the fitted physical laws to generate an augmented training dataset. The amount of data required for law fitting is significantly smaller compared with the traditional machine learning methods. Our previous study [8] shows that only a few seconds of microphone data were sufficient to fit the frequency response curve that governs the domain shift in a voice sensing task. Third, physics-informed machine learning improves the DNN model's performance. The conventional data augmentation applies *ad hoc* perturbations or transformations on the training data to improve the DNN models' robustness. However, such robustness will be lost if the distribution of the inference data is out of the scope of the augmented training data. Differently, physics-informed machine learning pinpoints the physical law governing the data generation or domain shifts and uses the fitted law to guide the DNN model learning. As shown in §3.2, the yielded DNN model exhibits better performance compared with the conventional approaches in our evaluation.

The contributions of this paper are summarized as follows:

- We identify the data scarcity and run-time domain shift challenges and their root causes in developing generalized machine learning algorithms for AIoT sensing. We propose to integrate the known physical laws in machine learning models to address these challenges.
- With a primary goal of encouraging more research in integrating physical laws into machine learning models in the field of AIoT sensing, we explicitly discuss the opportunities and challenges in this paper.
- We propose physics-informed class augmentation (PICA) and physics-informed domain adaptation (PIDA) approaches to address the unavailability of classes and domain shifts, respectively. The effectiveness of the proposed approaches is illustrated using three distinct case studies.

The rest of this paper is organized as follows. §2 reviews related work. §3 uses a set of case studies to demonstrate the opportunities of physics-informed machine learning. §4 discusses the challenges of physics-informed machine learning. §5 concludes this paper.

2 Related Work

■ **Self-supervised learning:** Despite the limited availability of labeled training data, AIoT sensing applications often have abundant unlabeled data. Recent advances in self-supervised learning inspire studies to apply it to improve the learning efficiency of AIoT sensing applications. LIMU-BERT [21] exploits the self-attention mechanism of the transformer to extract temporal relations from the unlabeled IMU sensor data. LIMU-BERT reduces the requirement of labeled data in IMU-based human activity recognition tasks. Contrastive learning is another effective feature representation learning technique. SimCLR [4] proposes an effective framework to construct positive/negative data pairs from unlabeled images. A feature extractor trained on such data pairs using contrastive loss can learn useful image representations without labels. Using the pre-trained model, SimCLR can achieve comparable results as supervised learning in image recognition tasks while requiring only 1% to 10% labeled training data. The studies mentioned above [4, 21] utilize machine learning techniques to extract useful feature representations from abundant unlabeled data. This paper discusses the possibilities of incorporating physical knowledge in the design of generalized machine learning models for AIoT sensing. Our approach has the potential to decrease the need for labeled data and result in more robust deep neural network models.

■ **Transfer learning:** Few-shot learning and generative adversarial learning-based domain adaptation are common transfer learning techniques for addressing domain shifts. MetaSense [6] employs the meta-learning technique to build a base model that can be adapted to a new target domain with a few data samples in voice sensing and human activity recognition tasks. Mic2mic [9] uses the unpaired source- and target-domain data to train a CycleGAN model for data translation. It is used to improve the voice sensing model's performance for an IoT microphone. In this paper, we consider the domain shifts governed by physical laws and use a small amount of source- and target-domain data to fit physical laws for DNN model transfer. Our proposed method identifies the cause of the domain shift and is more effective in addressing the run-time domain shifts in AIoT sensing applications.

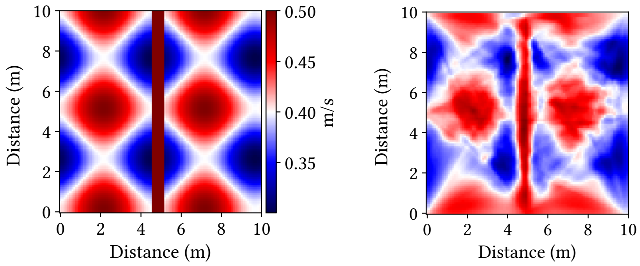
■ **Physics-informed machine learning** leverages the known prior knowledge to accelerate model training or improve the performance of machine learning algorithms. Its applications [13] include molecular properties prediction, fluid flow inference, edge plasma dynamics modeling, etc. According to a recent review [7], the current PIML approaches can be classified into three categories based on their representation of physics. The method based on *observational biases* utilizes the data generated by the physical laws to guide the training of DNN models. The DNN models trained on such data can capture the mechanism governing data generation. The method based on the *inductive biases* focuses on designing specialized neural network architectures with physical laws embedded. The study [18] provides an example of this method, in which the cooling and heating units in a data center space are incorporated into the architecture of the neural network for temperature prediction. The method based on *learning biases* incorporates physical laws into the DNN model's loss function. This yields a more efficient model training due to additional constraints imposed by the physical laws. The study [16] is an example of this method, in which the free-fall law is incorporated into the loss function of an object detection neural network. The methods based on *inductive biases* and *learning biases* require redesigns of neural network architecture and/or loss function. Differently, the method based on the *observational biases* does not modify the existing DNN designs and learning algorithms, rendering it universally applicable. Our approaches proposed in this paper belong to the method based on *observational biases*.

3 Opportunities

We use three case studies to illustrate the opportunities of physics-informed machine learning model generalization.

3.1 Physics-Informed Class Augmentation

In classification or regression tasks, a pre-trained DNN model will have poor performance on data samples from the classes unseen during the training phase. For example, a DNN model trained to recognize digits from 0 to 9 cannot determine the actual label if the tested data is a digit 10. A straw-man approach is to collect the data for new classes and retrain the DNN model. However,



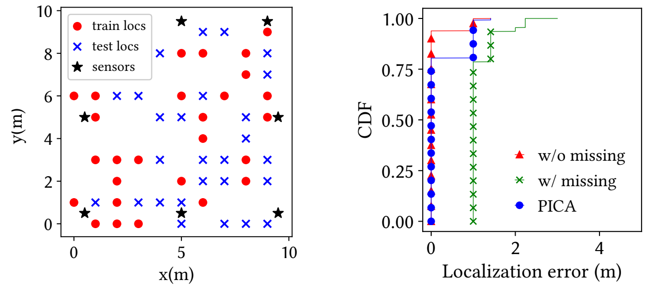
(a) True slowness model **(b) fitted slowness model**

Fig. 3: Floor slowness model fitting.

collecting training samples can result in unwanted overhead, due to the difficulty in separating the collection and labeling process of IoT data. PICA leverages physical laws that govern the data generation to augment training datasets for DNN model generalization. The procedures are as follows. First, we identify the parameterized physical model that governs the data generation for a specific AIoT application. The physical law may contain unknown parameters. Second, we use a small amount of genuine data to fit the unknown parameters of the physical model. Third, we use the fitted model to generate massive synthetic data for the unavailable classes or unavailable data points. Lastly, we use both the genuine and the augmented new class data to train a new recognition model that covers unknown classes. As a result, the new classification model can provide plausible predictions on unknown classes. We use a case study to demonstrate the advantages brought by PICA.

3.1.1 Case Study: Occupant localization at unknown locations. Occupant localization is a fundamental requirement of smart buildings. It can be used for crowd and energy management in museums or patient tracking in hospitals. Footstep-induced floor vibration [10] can be exploited for occupant localization. The floor vibration reaches sensors deployed on the floor at different times. If the signal propagation velocity is spatially consistent, the time differences of arrival (TDoA) can be calculated to determine the source of the signal using triangulation. However, the uneven floor medium causes the vibration wave propagates at a non-constant speed across the floor, leading to the inaccurate result of the TDoA approach. Thus, the fingerprinting approach is employed to collect footsteps' fingerprints and train DNN models to recognize footsteps' locations. We investigate the challenge of unavailable class issues and the benefit brought by the proposed PICA for a footstep localization system. This case study is similar to our published work [8] while bearing a different objective. Our earlier work [8] focuses on adapting an event source localization model from a homogeneous medium to a heterogeneous medium, which is a domain adaptation problem. Our current study focuses on improving the generalization of a DNN model to unseen locations in a heterogeneous medium.

We simulate footstep localization on a 10m $\times$ 10m floor with an uneven medium. The wave propagation in the floor is depicted by the floor's *slowness model*, which is denoted as $\boldsymbol{s}$. The floor consists of 100×100 grids, each grid is associated with a distinct wave propagation speed as shown in Fig. 3a. The color bar value corresponds to the wave propagation speed (m/s). The footstep-induced vibration signal propagates through the uneven medium and is detected by 8 sensors placed on the floor. Fig. 4a shows the



(a) Footsteps locations **(b) Localization performance**

Fig. 4: PICA for occupant localization.

simulated footsteps and deployed sensors, where dots and crosses represent the footstep locations and stars represent the sensor locations. A total of 1000 footstep events are triggered at each location with a 20cm radius. A 5% random noise is added to each sensor's reading to introduce bias. The TDoA is calculated as the fingerprint of a footstep based on a reference sensor.

Collecting fingerprints with ground-truth locations is difficult in a building. Data collection is typically performed through manually triggered events at a limited number of spots, e.g., hiring volunteers to stimulate step vibration at marked locations. As a result, the fingerprints only cover a small portion of the floor, leading to unavailable classes and data points in the training dataset. This results in a degraded performance of the pre-trained model when applied to data collected outside of the fingerprinted locations. To illustrate this phenomenon, we train a location recognition model using the fingerprints collected at locations marked by the red dots in Fig. 4a. The plot labeled "w/o missing" in Fig. 4b shows the cumulative distribution function (CDF) of the localization errors. The mean localization error is 0.05 m. When the model is tested using data collected at locations marked by blue crosses, which are not part of the training dataset, the plot labeled "w/ missing" shows the result. The mean localization error increases to 1.8 m, indicating a significant performance degradation of the classification DNN model on data from unavailable classes.

PICA improves the DNN model's performance on unavailable classes through four steps. First, we identify that the physical law governing the wave propagation is the floor's slowness model $\boldsymbol{s}$, which contains the unknown propagation speed value for each grid. Second, we use a small amount of real data to estimate the propagation speed of $\boldsymbol{s}$ based on the Bayesian Algebraic Reconstruction Technique algorithm [20]. Third, we use the estimated slowness model $\hat{\boldsymbol{s}}$ to generate a large amount of synthetic data that can cover the locations marked by the crosses. Fig. 3b shows the fitted slowness model using 100 footstep data, which closely resembles the ground truth. Lastly, we train a new location recognition model using both synthetic and real training data. The plot labeled "PICA" shows the result when the new model is evaluated on the data from locations marked by the crosses. The mean error decreases to 0.15 m. PICA effectively tackles the unavailable class problem by augmenting the data with fitted physical laws.

3.2 Physics-Informed Domain Adaptation

A sensing DNN model pre-trained on the standard dataset suffers a performance drop after deployment due to the run-time domain

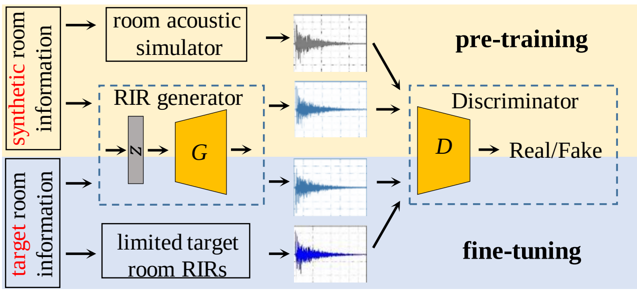


Fig. 5: Training procedures of RIR generator.

shift. PIDA considers the applications where the domain shifts are governed by certain physical laws. It applies the following four steps to address the data scarcity problem in the target domain. First, we identify the physical law governing the domain shift for a specific application, which contains unknown parameters. Second, we use a small amount of real data collected in both the source and target domain to fit the physical law. Third, we use the fitted law to augment the source-domain data to the target domain. Lastly, we transfer the DNN model to the target domain using the augmented dataset, through re-training or fine-tuning.

We use two voice-sensing case studies to demonstrate the benefits of PIDA. The case studies differ in terms of the complexity of the physical laws. In the first case study where the physical law can be parameterized, we use a small amount of real data from the target domain to estimate the unknown parameters. In the second case study where the physical law cannot be parameterized, we train a neural network as the surrogate model of the physical law and use it to generate augmented training data for model transfer.

3.2.1 Case study 1: Addressing microphone heterogeneity for speech recognition. Human voice recognition is a key component of smart building technology. DNN-powered virtual assistants such as Amazon Echo and Google Nest have become an integral part of our daily routines, providing smooth human-machine interaction. However, the commonly used low-cost IoT microphones for voice sensing have varying quality and can result in domain shifts between recorded data and the standard datasets used for model training. This can affect the accuracy of pre-trained DNN models. We identify the domain shift is caused by the microphone's frequency response curve (FRC) and propose a microphone profiling method to estimate different microphones' FRCs. Then, we apply PIDA to generate augmented target microphone data for model adaptation.

In the experiment where the microphone hardware heterogeneity causes word error rate (WER) increase for a pre-trained automated speech recognition model, PIDA can reduce the WER increase from 50% to 70% on evaluated microphones. This result outperforms the baseline approaches that are based on conventional data augmentation or data calibration. A detailed evaluation can be found in our previous published work [8].

3.2.2 Case study 2: Adapting audio sensing model in new environments. In addition to the device hardware heterogeneity, a voice-sensing device's ambient environment can affect its voice recognition performance. The impact of the environment is a result of the room's acoustic response to the sound source, which is measured

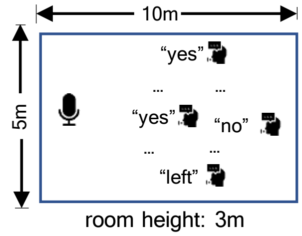


Fig. 6: Voice data collection.

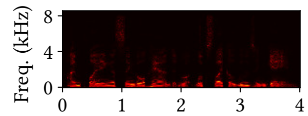


Fig. 7: Original voice.

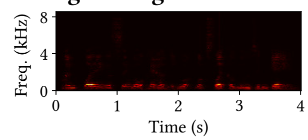
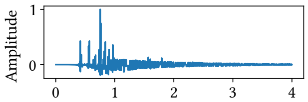


Fig. 8: Received voice.

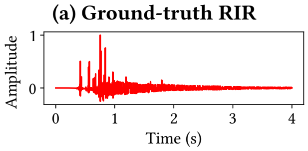
by the room impulse response (RIR). Thus, each deployment environment forms a target domain. In practice, RIRs can be collected to adapt a standard voice-sensing model to a target room. However, the collection of RIRs requires specialized equipment and can involve significant manual effort. As a result, only a limited number of public RIRs datasets [2] are available. Several acoustic simulators have been created to produce realistic RIRs based on room parameters. However, the simulators require pre-defined parameters and exceptional computational resources for RIRs generation. They are not configurable for a real target room. Therefore, we propose using a DNN model to generate RIRs in a target room. The DNN model acts as a substitute for the physical law governing the RIRs generation. The DNN model also provides the flexibility to be configured for a target room if a small amount of data are given. In what follows, we apply this approach to adapt a keyword spotting (KWS) model to new environments.

We train a KWS model using Google Speech Commands dataset [19] to recognize 10 keywords. The pre-trained model achieves an oracle recognition accuracy of 90% on the standard testing dataset. We collect new voice command data in a simulated room to evaluate the KWS model's performance. Fig. 6 shows the data collection process. We deploy a microphone near the left wall of a 5 m×10 m×3 m room to record the voice commands triggered at random locations within the room. The received microphone data is a convolutional result of the voice command and the room's RIR at a specific location. Fig. 7 and Fig. 8 show the spectrograms of the original and the recorded voice command. There are salient differences between the data samples. When we apply the pre-trained KWS model to the recorded voice commands. The bar labeled "Unmodified" in Fig. 9c shows the result, where the recognition accuracy drops to 73%. Thus, the data collected in the room form a new target domain and the voice-sensing DNN model needs to be adapted.

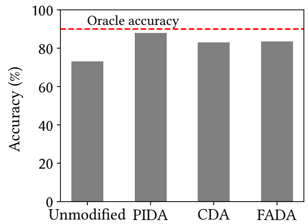
PIDA for KWS model adaptation in a room consists of four steps. First, we identify that the domain shift is governed by a room's RIRs. As the conventional acoustic simulators lack reconfigurability, we opt to use a DNN to model the physical law for RIRs generation. We adopted the adversarial training methods described in [14] to train a DNN-based RIR generator with the guidance of the room acoustic simulator. As illustrated in the upper part of Fig. 5, given a set of parameters, e.g., room dimension, speaker and microphone location, both the room acoustic simulator and the RIR generator network generate RIR samples that are then discriminated by a discriminator network. At end of the adversarial training process, the RIR generator is capable of producing RIRs that closely resemble those produced by the room acoustic simulator. Third, as shown in the lower part of Fig. 5, we use RIRs collected at limited locations



(a) Ground-truth RIR



(b) Generated RIR



(c) KWS model accuracies of different approaches

Fig. 9: Adaptating a KWS model to a new environment (The target room is simulated with Pyroomacoustics [15]).

from the target room to finetune the RIR generator. As such, the finetuned RIR generator is used to produce augmented RIRs at unknown locations in the target room. Fig. 9a and Fig. 9b show the ground truth and the generated RIRs in a simulated target room. The data samples exhibit a cosine similarity of 0.9. Lastly, we use the finetuned RIR generator to produce large amounts of synthetic RIRs in the target room for KWS model adaptation. The bar labeled "PIDA" shows the new result, the recognition accuracy increases to 87%. It outperforms the baseline approaches that are based on conventional data augmentation (CDA) and few-shot adversarial domain adaptation (FADA) [11]. Their recognition accuracies are 82.5% and 83%, respectively.

4 Challenges

We discuss several challenges when applying physics-informed machine learning model generalization in AIoT sensing applications.

Identifying the physical laws is challenging. The proposed PICA and PIDA rely on the identified physical knowledge to generate augmented data for DNN model generalization. However, in many AIoT sensing applications, the sensed data is often governed by intricate physical processes, making it difficult to determine the physical laws. When physical laws are elusive, we may turn to traditional data-driven methods to train DNN models.

The amount of data needed for physical law fitting remains unresolved. The ability to accurately fit physical laws is crucial for the generalization of physics-informed machine learning models. The effectiveness of the models depends on the quality of the augmented data generated by the fitted laws. Unfortunately, there is currently no theoretical analysis available to guide the determination of the required data for fitting physical laws. Intuitively, systems with simple physical laws require less data, while systems with complex physical laws will require more data for law fitting.

The evaluation of systems with multiple physical laws is lacking. The used three case studies consider the system governed by a single physical law. In many AIoT sensing applications, the data generation or domain shifts can be governed by multiple physical laws. A possible approach is through the separation of physical laws. However, it may require additional genuine data to fit the separated laws.

Considering the potential benefits brought by physics-informed machine learning generalization, the above challenges should not discourage us from pursuing more efficient approaches in incorporating physical knowledge in machine learning algorithms for AIoT sensing applications.

5 Conclusion

This paper discusses the potential of generalizing physics-informed machine learning models in AIoT sensing applications. We propose *physics-informed class augmentation* and *physics-informed domain adaptation* to address the scarcity of labeled training data and runtime domain shifts problems. The benefits of the physics-informed machine learning model generalization are illustrated using three distinct case studies. We also discuss the challenges when applying these methods to specific AIoT applications. We aim to spark more investigation into utilizing physical knowledge to enhance machine learning algorithms in cyber-physical systems.

Acknowledgments

This research is supported by the Ministry of Education, Singapore, under its Academic Research Fund Tier 1 (RG88/22).

References

[1] [n.d.]. https://www.mturk.com/.
[2] [n.d.]. https://github.com/RoyJames/room-impulse-responses.
[3] Ting Chen, Simon Kornblith, Mohammad Norouzi, and Geoffrey Hinton. 2020. A simple framework for contrastive learning of visual representations. In *ICML*. PMLR, 1597–1607.
[4] Ting Chen, Xiaohua Zhai, Marvin Ritter, Mario Lucic, and Neil Houlsby. 2019. Self-supervised gans via auxiliary rotation loss. In *CVPR*. 12154–12163.
[5] Jia Deng, Wei Dong, Richard Socher, Li-Jia Li, Kai Li, and Li Fei-Fei. 2009. Imagenet: A large-scale hierarchical image database. In *CVPR*. IEEE, 248–255.
[6] Taesik Gong, Yeonsu Kim, Jinwoo Shin, and Sung-Ju Lee. 2019. MetaSense: few-shot adaptation to untrained conditions in deep mobile sensing. In *Proceedings of the 17th Conference on Embedded Networked Sensor Systems (SenSys)*. 110–123.
[7] George Em Karniadakis, Ioannis G Kevrekidis, Lu Lu, Paris Perdikaris, Sifan Wang, and Liu Yang. 2021. Physics-informed machine learning. *Nature Reviews Physics* (2021), 422–440.
[8] Wenjie Luo, Zhenyu Yan, Qun Song, and Rui Tan. 2021. Phyaug: Physics-directed data augmentation for deep sensing model transfer in cyber-physical systems. In *IPSN*. 31–46.
[9] Akhil Mathur, Anton Isopoussu, Fahim Kawsar, Nadia Berthouze, and Nicholas D Lane. 2019. Mic2mic: using cycle-consistent generative adversarial networks to overcome microphone variability in speech systems. In *IPSN*. 169–180.
[10] Mostafa Mirshekari, Shijia Pan, Jonathon Fagert, Eve M Schooler, Pei Zhang, and Hae Young Noh. 2018. Occupant localization using footstep-induced structural vibration. *Mechanical Systems and Signal Processing* (2018), 77–97.
[11] Saeid Motiian, Quinn Jones, Seyed Iranmanesh, and Gianfranco Doretto. 2017. Few-shot adversarial domain adaptation. *NIPS* 30 (2017).
[12] Vassil Panayotov, Guoguo Chen, Daniel Povey, and Sanjeev Khudanpur. 2015. Librispeech: an asr corpus based on public domain audio books. In *ICASSP*. IEEE.
[13] Maziar Raissi, Paris Perdikaris, and George E Karniadakis. 2019. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *Journal of Computational physics* (2019).
[14] Anton Ratnarajah, Shi-Xiong Zhang, Meng Yu, Zhenyu Tang, Dinesh Manocha, and Dong Yu. 2022. FAST-RIR: Fast neural diffuse room impulse response generator. In *ICASSP*. IEEE, 571–575.
[15] Robin Scheibler, Eric Bezzam, and Ivan Dokmanić. 2018. Pyroomacoustics: A python package for audio room simulation and array processing algorithms. In *2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 351–355.
[16] Russell Stewart and Stefano Ermon. 2017. Label-free supervision of neural networks with physics and domain knowledge. In *AAAI*, Vol. 31.
[17] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Łukasz Kaiser, and Illia Polosukhin. 2017. Attention is all you need. *NIPS* 30 (2017).
[18] Ruihang Wang, Xin Zhou, Linsen Dong, Yonggang Wen, Rui Tan, Li Chen, Guan Wang, and Feng Zeng. 2020. Kalibre: Knowledge-based neural surrogate model calibration for data center digital twins. In *BuildSys*. 200–209.
[19] Pete Warden. 2018. Speech commands: A dataset for limited-vocabulary speech recognition. *arXiv preprint arXiv:1804.03209* (2018).
[20] Stephen Wright, Jorge Nocedal, et al. 1999. Numerical optimization. *Springer Science* 35, 67-68 (1999), 7.
[21] Huatao Xu, Pengfei Zhou, Rui Tan, Mo Li, and Guobin Shen. 2021. Limu-bert: Unleashing the potential of unlabeled data for imu sensing applications. In *SenSys*.