A Survey of Few-Shot Learning for Radio Frequency Fingerprint Identification

Hao Li[1], Yu Tang[1(✉)], Di Lin[1], Yuan Gao[2], and Jiang Cao[2]

[1] University of Electronic Science and Technology of China, Chengdu, Sichuan, China
yutang@uestc.edu.cn
[2] Military Academy of Sciences, Beijing, China

Abstract. With the development of the Internet of Things technology, the radio frequency (RF) fingerprint identification technology of wireless communication equipment has also risen, providing new ideas for network security and RF perception systems. The existing RF fingerprint identification technology is mainly based on traditional machine learning or deep learning. In the face of small sample data or data imbalance, the classification effect is not satisfactory. Therefore, in this paper, we propose the use of Few-Shot Learning (FSL) to solve the problem of radio frequency fingerprint small sample recognition. We review the current RF fingerprint identification technology and FSL methods. What's more, we analyze some available methods from two aspects. (i) From the perspective of data, the samples of RF signal training data set can be expanded manually or by using transformation function, can also be generated by generative model; (ii) From the perspective of algorithms, prior knowledge can be used to train the new model through fine-tuning, metric, and meta-learning. Finally, we look forward to the challenges and opportunities that the RF fingerprint identification technology may face from theory and application.

Keywords: Radio frequency fingerprint · Few-shot learning · Meta-learning

1 Introduction

In recent years, with the rapid development of the Internet of Things and artificial intelligence technology, wireless communication devices have played an irreplaceable role in the civilian and military fields, and how to ensure the security of wireless communication devices in the network is particularly important. Compared with wired network, wireless network is more vulnerable to attack, because traditional methods to ensure wireless network security are usually based on cryptographic mechanisms and security protocols, and wireless communication devices exposed to the network still have security risks. Therefore, people urgently need a new type of security mechanism to effectively identify authorized users and unauthorized users to reduce potential threats in the network.

Different wireless communication devices emit different signals due to hardware differences. The hardware feature extracted by analyzing the small difference of the

X. Wang et al. (Eds.): AICON 2021, LNICST 396, pp. 433–443, 2021.
https://doi.org/10.1007/978-3-030-90196-7_37

radio frequency (RF) signal is called the Radio Frequency Fingerprint [1] of the device, and the method of using RF fingerprint to identify different wireless communication devices is called RF fingerprint identification.

At present, the existing RF fingerprint identification technology mainly adopts traditional machine learning or deep learning to train a large number of signal data sets of a group of wireless communication devices. After extracting signal features as RF fingerprint, these methods construct a classifier based on the training model and training parameters to confirm the identity of the wireless communication device. However, in some cases, we can't obtain a large number of I/Q signal samples, but a small amount of data can be collected. The signal of these communication equipment, such as shortwave communication equipment for long-distance communication or short-range tactical communication, medium-wave communication equipment for emergency communications, with relatively concealed environment is weak and has long transmission period, which is not easy for the receiver to collect. It cannot obtain a large amount of data, resulting in the signal sample data imbalance. In the case of small data sets, traditional machine learning and deep learning approaches for data-intensive applications are no longer applicable because too few training sets are prone to reduce the accuracy of the algorithm and overfitting.

Recently, a machine learning method, few-shot learning (FSL) is proposed. Using prior knowledge, FSL can quickly generalize to new tasks that contain only a small number of samples with supervised information. Therefore, it can be considered to combine FSL with RF fingerprint, in order to solve the small sample problem of RF fingerprint.

The contributions of this survey are as follows:

- We discuss how to perform RF fingerprint identification through FSL in the case of insufficient or unbalanced data sources to help confirm the identity of wireless communication devices.
- According to the existing FSL methods, two types of RF fingerprint small sample identification methods are proposed: On the one hand, from the perspective of data, the data volume of RF signal is expanded by means of manual or Generative Adversarial Network (GAN). On the other hand, from the perspective of algorithm, the prior knowledge is used to fine-tune the existing RF fingerprint training model parameters.
- We look forward to the development trends and application prospects of RF fingerprint small sample identification technology.

The remainder of this survey is organized as follows. Section 2 provides overview for RF fingerprint identification and FSL. Section 3 is for methods that augment data to deal with RF fingerprint small sample identification problem. Section 4 is for methods that optimize the algorithm to solve RF fingerprint small sample identification problem. Finally, the survey closes with conclusion in Sect. 5 And we propose future directions for RF fingerprint small sample identification in terms of theories and applications.

2 Overview

At present, RF fingerprint technology is becoming more and more developed. And FSL has also been paid attention to and applied in the field of image and text recognition. This section mainly reviews the research status of RF fingerprint identification technology based on traditional machine learning in recent years, and introduces FSL.

2.1 RF Fingerprint Identification Technology

As shown in Fig. 1, RF fingerprint identification usually includes three stages: signal collection, pretreatment and RF fingerprint extraction. According to the requirements of RF fingerprint, the identification system needs to carry out several preprocessing processes for the collected signals, such as phase compensation, energy normalization and discarded unqualified signal. The fingerprint extraction mainly includes training and classification stage. In the training stage, the host receiver receives the signal from the device for sampling, and then extracts the features to generate the RF fingerprint. In the classification stage, after receiving signals from the devices to be identified, the receiver will classify the devices according to the similarity of these features by comparing the characteristics of each type.

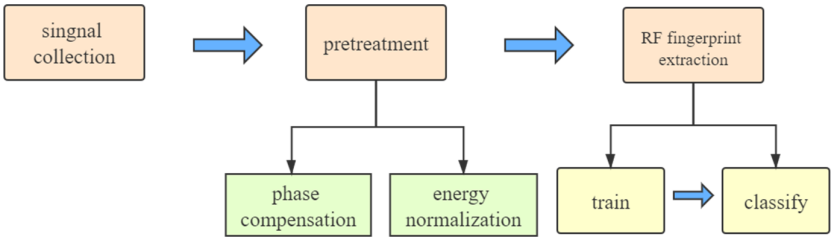

Fig. 1. RF fingerprint identification process.

Traditional Machine Learning for RF Fingerprint Identification. The traditional machine learning can identify the transmitter by extracting the signal features and comparing fingerprint database. In [2], a novel approach has been proposed to use a digital noise radar (DNR) to actively interrogate microwave devices and classify defective units by using radio frequency distinct native attribute (RF-DNA) fingerprinting and various classifier algorithms, such as multiple discriminant analysis/maximum likelihood (MDA/ML) and generalised relevance learning vector quantisation-improved (GRLVQI) classifiers. In [3] , the feasibility of RF-DNA extraction from different devices is successfully proved by using MDA/ML classifier, and one-to-many device classification and one-to-one device ID verification are performed by using Gabor-based method. It is proved that using dimensionality reduction analysis can reduce the number of required fingerprint features while maintaining consistent discrimination performance. In [4], the RF fingerprinting is proposed by extracting the parameter characteristics such as

information dimension, constellation feature and phase noise spectrum in the transmitted information when it is applied to the universal software radio peripheral (USRP) software defined radio (SDR) platform. In proposed methods, the traditional support vector machine (SVM) classifier, the machine-based integrated classifier bagged tree and the adaptive weighting algorithm weighted k-nearest neighbor (KNN) achieved good classification performance under different signal-to-noise ratios (SNR).

Deep Learning for RF Fingerprint Identification. Deep learning relies on large amounts of data to train the models that can recognize existing RF fingerprint without the need to extract data features in advance, as shown in Fig. 2. In [5], a long short-term memory (LSTM) based recurrent neural network is proposed and used for automatically identifying hardware-specific features and classifying transmitters. In [6], By combining the software defined radio (SDR) sensing capability with the convolutional neural network (CNN) morphing on the basis of AlexNet, a specific radio frequency can be uniquely identified between similar devices. The main advantage of this approach is that the CNN learning framework operates on the raw I/Q sample. In [7], data diversity in the training set is key to extracting broadly-applicable features and achieving high accuracy with incremental models. The implementation of the above methods is based on a large amount of data in a specific environment but is not applicable to small sample data sets. Therefore, we need find new method to solve this problem.

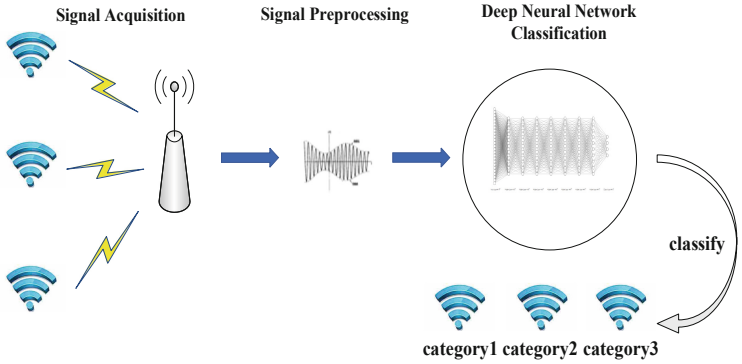

Fig. 2. Deep learning based RF fingerprint identification framework.

2.2 Few-Shot Learning

Few-Shot Learning (FSL) is a machine learning method proposed in recent years to solve the problem of small amount of data and data imbalance. It makes use of prior knowledge to quickly generalize to a new task containing only a small number of samples with supervised information, so as to help relieve the burden of collecting large-scale supervised data. Transfer Learning [8] and Meta-Learning [9] methods transfer prior knowledge from source task to small sample task, which is widely used in FSL. In [10], Yaqing Wang et al. point out that the core issue of FSL supervised learning problem is

the unreliable empirical risk minimizer. According to the utilization of prior knowledge, FLS methods can be divided into three categories: data, model and algorithm. From the perspective of data, signal samples with less data can be expanded, such as transforming samples based on training data, transforming samples based on weakly labeled or unlabeled data, transforming samples based on similar data sets, etc. From the perspective of model, multi-task learning, embedded learning, learning based on external storage and generative modeling can be carried out. From the perspective of algorithm, the existing training model parameters can be fine-tuned and the optimizer can be learned.

Therefore, FSL provides a new idea for solving the imbalance problem of small sample identification data of RF fingerprint and improving the generalization ability of the model. This paper will demonstrate the feasibility of FSL in RF fingerprint identification technology from two aspects of augmenting data and optimizing algorithm.

3 Data Augmentation

When the amount of collected RF signal data is insufficient, the RF signal data set can be augmented in different ways to obtain enough samples, and then the machine learning or deep learning algorithm mentioned in Sect. 2 can be used for RF fingerprint identification. Table 1 shows the classification of methods used to enrich RF signal samples.

Table 1. Data augmentation methods for RF sample. The transformer t takes input (x, y) and returns generated sample (x', y') to augment the few-shot RF sample data.

Method	Input (x, y)	Transformer t	Output (x', y')
Augmenting samples from RF training data sets	Raw data (x, y)	Artificial augmentation or a transformation function	$(t(x), y)$
Generating samples by generative model	Fabricating data (x, y)	A generator to generate new data sets	$(t(x), t(y))$

3.1 Augmenting Samples from RF Training Data Set

The strategy expands the data set by transforming each RF signal sample into several changing samples with transformation. We can do manual amplification, for example by using MATLAB and other tools for signal simulation to get more data. Besides, after the signal is transformed in the time-frequency domain to get the spectrum image, the data is expanded by flipping, translation, scaling, rotation and other methods. However, it requires expensive labor cost and depends heavily on the knowledge of communication domain.

An earlier FSL learned a set of geometric transformations, which is applied to each sample, from similar classes by iteratively aligning each sample with other samples. Similarly, in [11], a single transformation function is learned to transfer the variation

between sample pairs learned from other categories to the sample pair. Therefore, we can consider constructing a transformation function that automatically expands the data set based on the existing RF signals to form a large data set that can then be learned by standard machine learning methods, as shown in Fig. 3.

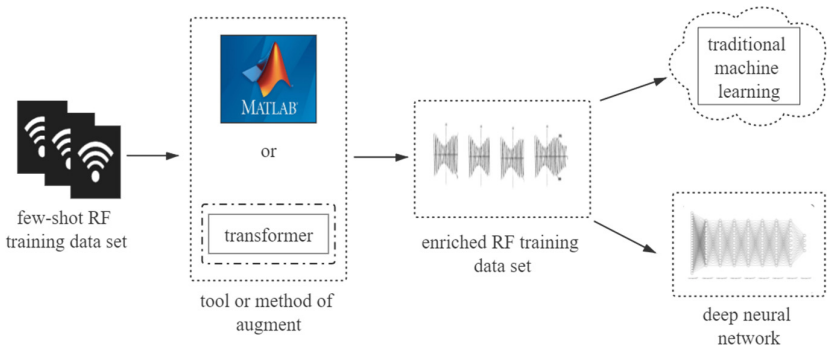

Fig. 3. Data augmentation methods for few-shot RF training data set.

3.2 Generating Samples by Generative Model

The strategy expends the data set via generative model. In essence, the generative model is a kind of maximum likelihood estimation, which is used to generate the model of the specified distribution data. The distribution generated by the generator we have now can be assumed to be $P_G(x;\theta)$, and this is a distribution controlled by θ, where θ is the parameter of the distribution. Suppose we take some data out of the real distribution, $\{x_1, x_2 \ldots, x_m\}$, and then we want to compute a likelihood $P_G(x_i;\theta)$, as shown in Formula (1).

$$L = \prod_{i=1}^{m} P_G(x_i;\theta) \tag{1}$$

In order to maximize this likelihood, we need to find a θ^*.

$$\theta^* = \underset{\theta}{arg\,max} \prod_{i=1}^{m} P_G(x_i;\theta) \tag{2}$$

Therefore, a generative adversarial network (GAN) [12] is designed to generate indistinguishable synthetic samples. GAN has two networks, one generator that is responsible for fabricating data and one discriminator that is responsible for judging the authenticity of data. The best generation effect is achieved by two networks against each other. In order to augment data, we can use the existing few-shot RF samples to establish the generation model via GAN.

In Fig. 4, the generator can fabricate fake RF samples based on noise samples, and use these fake samples to deceive the discriminator. The discriminator is responsible for identifying whether this is a real or fake sample and will give a score. When the score

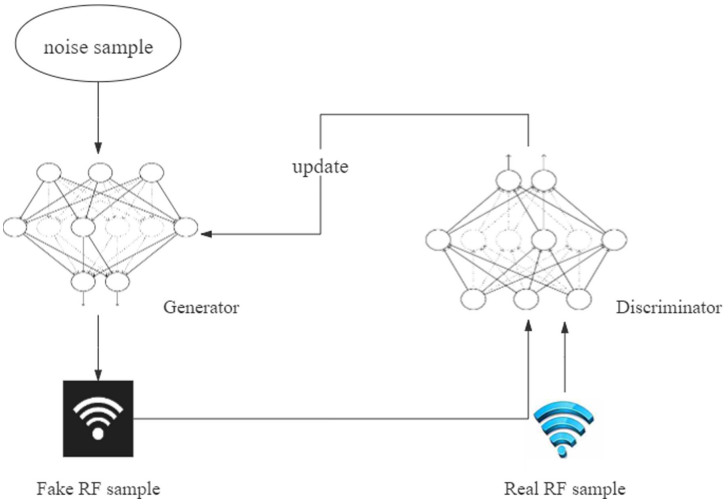

Fig. 4. A GAN model for generating RF samples.

value is high, it indicates that the discriminator can effectively distinguish between true and false samples, but the effect of the generator is not good, so the parameters of the generator needs to be adjusted. However, generative modeling methods have high inference cost, and are more difficult to derive than deterministic models.

3.3 Summary

The choice of which method to expand data depends on the application scenario. When there are only a small number of RF signal samples, the method of augmentation can be manually generated or transformed function. When there are a small number of target RF signal samples and other data sets similar to the sample, GAN can be used for data augmentation. Then standard machine learning methods are used for RF fingerprint identification. However, it is only suit for specific data with specific rules, but not easy to use on other data sets. Therefore, in the case of not changing the existing RF samples, we can also consider from the perspective of model algorithm.

4 Optimization of Algorithm

Due to the limitation of data, the problem of small sample identification of RF fingerprint can also be solved from the perspective of algorithm. No matter what model, it needs the support of data to carry on the training. We can use the prior knowledge to adjust and optimize the existing model parameters and apply it to the small sample data set of RF signals. This section presents several learning methods, as shown in Table 2.

4.1 Finetune

This method has been widely used in Transfer Learning. A certain amount of annotated data is obtained and then fine-tuned based on a pre-trained model. This pre-trained model

Table 2. Optimization methods of Algorithm for few-shot RF samples.

Method	Supervised information	Prior knowledge
Finetune	A few labeled RF samples for target class	Pre-trained basic model
Metric	A few labeled RF samples for each class	Embedding learned model
Meta-learning	A few labeled RF samples for each class of the target task	Meta learner

is obtained from large datasets with rich tags, such as the Oracle RF Fingerprinting Dataset, known as the Common Data Domain. Then the training is carried out on a small sample data domain of some RF signal. During training, the parameters of the pre-trained model will be fixed and specific model parameters will be trained. There are many training tricks, including how to set the fixed layer and learning rate, etc. This method can be relatively fast and does not have to rely on too much data.

4.2 Metric

The method is to model the distance distribution between samples, so that the samples the same kind are close, and the samples of different kinds are far away. By learning an end-to-end nearest neighbor classifier, it benefits from the advantages of both parameters and no parameters, in order that it cannot only learn new samples quickly, but also have good generalization for known samples. Siamese Networks [13] judges whether they belong to the same class through the distance of sample pairs. Then the features extracted from the network are reused for One/Few-Shot Learning. In Prototypical Network [14], each category has a prototype representation whose archetype is the mean value of the support set in the embedding space. The classification problem then becomes a problem of the nearest neighbor in the embedding space. Matching Networks [15] builds different encoders for the support set and the batch set. The output of the final classifier is the weighted sum of predicted values between support set and query set, which can generate labels for unknown categories under the premise of not changing the network model. Therefore, the above three network models can be considered to apply to the identification of few-shot RF fingerprint.

4.3 Meta-learning

Meta-Learning acquires meta-knowledge through meta-training on known tasks, which helps model to learn quickly on new tasks with a good ability of model generalization. In Fig. 5, meta-learning task is divided into training and testing in two stages.

As for parameter θ in the Fig. 5, At the t th iteration, $\theta_t = \theta_{t-1} + \Delta\theta_{t-1}$, where $\Delta\theta_{t-1}$ is the update. For the popular stochastic gradient descent (SGD), θ is updated as

$$\theta_t = \theta_{t-1} - \alpha_t \nabla_{\theta_{t-1}} \ell(h(x_t; \theta_{t-1}), y_t), \tag{3}$$

where αt is the step size. With θ initialized at θ_0, θ_t can be written as

$$\theta_t = \theta_0 + \sum_{i=1}^{t} \Delta\theta_{t-1}. \tag{4}$$

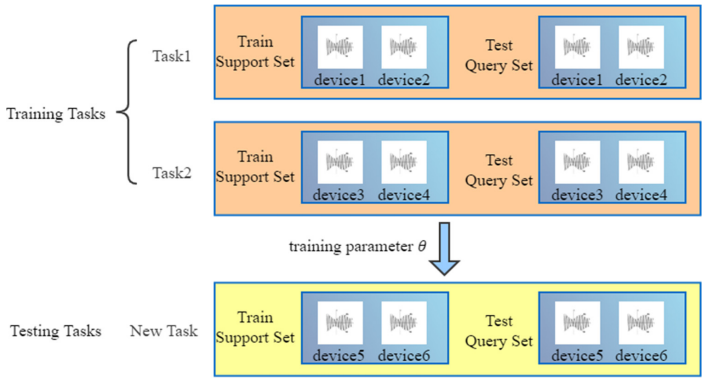

Fig. 5. The two stage of meta-learning task. In the first stage, training a large number of tasks to get a training parameter θ. In the second stage, in the face of a new category of tasks, there is no need to change the existing model to classify by updating parameter θ.

MAML [16] is an example of an approach that learns from prior knowledge and provides a good initialization parameter. The MAML algorithm is independent on model, which means it can be applied to any model, with the only requirement being that the model is trained using gradient descent. The idea is to learn an initialization parameter that can achieve good results with only a few steps of gradient descent using a small number of samples when a new task is encountered. In [17], A Meta Learner model based on LSTM is proposed for learning optimization of algorithm, which is used to train another learner neural network classifier in the case of small samples. The model learns a general initialization for the learner network, the classifier network, which can help the learner network to converge quickly during training. Similarly, the above two Meta-Learning methods can be adjusted and then applied to the identification of few-shot RF fingerprint.

4.4 Summary

The fine-tune method can help identify few-shot RF fingerprint via the existing pre-trained model parameters, but this strategy may sacrifice precision for speed. Metric method can be used to train a network model by calculating the spacing of few-shot RF samples without existing parameter, which has good generalization for known samples. Meta-Learning method mainly obtains a learner through multi-task training, which can use fewer RF samples to learn to identify new RF fingerprint. All of these methods can be used as solutions for small sample identification of RF fingerprint.

5 Conclusion and Future Work

RF fingerprint identification technology is a non-cryptographic security authentication technology. It can be identified by using the hardware difference of the device, which effectively solves the security hidden danger and controls the power consumption. The combination of small sample learning can help reduce the burden of collecting large

scale supervised RF signal data in applications. In this survey, we review the recent research status of FSL and RF fingerprint identification technology based on machine learning. Then, from the point of view of data and algorithm, the FSL methods that can be used in RF fingerprint small sample identification are analyzed and summarized. Finally, we will discusse the challenges and opportunities faced by FSL in RF fingerprint identification technology from the perspective of theories as well as the problems to be solved and studied urgently, and puts forward the future application scenarios of RF fingerprint identification.

5.1 Theories

At present, small sample identification technology for RF fingerprint is still in the exploration and research stage. Traditional machine learning methods are only aimed at the processing of massive data, while FSL is only studied in the field of image and text. Therefore, we can try to use the methods mentioned in Sect. 3 and Sect. 4 to construct the RF fingerprint small sample identification model from two perspectives of data or algorithm. It can solve the overfitting problem and improve the generalization ability of the existing model with few-shot RF sample or imbalance of data.

5.2 Applications

RF fingerprint technology as a kind of high reliability of wireless access technology, does not need to send additional access authentication information. Under the background of a variety of applications have some practical.

The small sample identification technology of RF fingerprint can be applied to the wireless Internet of Vehicles under the background of 5G. The use of radio frequency fingerprints to replace traditional license plate vehicle identity authentication methods not only saves the cost of redundant information transmission, but also has the advantages of non-tampering, uniqueness, and security. Using the RF fingerprint of the terminal device as the authentication information of the access network can effectively solve the problem of access power consumption. Using RF fingerprints as an indicator of security authentication can effectively solve the multi-port security access problem under the 5G background. In the future, RF fingerprint identification technology will enter people's lives and improve people's quality of life.

Acknowledgement. Partially Funded by Science and Technology Program of Sichuan Province (2021YFG0330), partially funded by Grant SCITLAB-0001 of Intelligent Terminal Key Laboratory of SiChuan Province, and partially Funded by Fundamental Research Funds for the Central Universities (ZYGX2019J076)).

References

1. Polak, A.C., Goeckel, D.L.: Identification of wireless devices of users who actively fake their RF fingerprints with artificial data distortion. IEEE Trans. Wireless Commun. **14**(11), 5889–5899 (2015)

2. Lukacs, M., Collins, P.: Classification performance using 'RF-DNA' fingerprinting of ultra-wideband noise waveforms. Electron. Lett. **51**(10), 787–789 (2015)
3. Reising, D.R., Temple, M.A., Jackson, J.A.: Authorized and rogue device discrimination using dimensionally reduced RF-DNA fingerprints. IEEE Trans. Inf. Forensics Secur. **10**(6), 1180–1192 (2015)
4. Hu, S., Lin, D.: Machine learning for RF fingerprinting extraction and identification of soft defined radio devices. Lecture Notes in Electrical Engineering 572, 189–204 (2020)
5. Wu, Q., Feres, C., Kuzmenko, D.: Deep learning based RF fingerprinting for device identification and wireless security. Electron. Lett. **54**(24), 1405–1407 (2018)
6. Riyaz, S., Sankhe, K., Ioannidis, S.: Deep learning convolutional neural networks for radio identification. IEEE Commun. Mag. **56**(9), 146–152 (2018)
7. Youssef, K., Bouchard, L.S., Haigh, K.Z.: Machine learning approach to RF transmitter identification. IEEE J. Radio Freq. Identification **2**(4), 197–205 (2018)
8. Wu, L., Wang, Y., Yin, H.: Few-shot deep adversarial learning for video-based person re-identification. IEEE Trans. Image Process. **29**, 1233–1245 (2020)
9. Santoro, A., Sergey, B.: Meta-Learning with memory-augmented neural networks. In: International Conference on Machine Learning, ICML 4, pp. 2740–2751 (2016)
10. Wang, Y., Yao, Q., Kwok, J.T.: Generalizing from a few examples: a survey on Few-shot Learning. ACM Comput. Surv. **53**(3), 1–34 (2020)
11. Hariharan, B.: Low-shot visual recognition by shrinking and hallucinating features. Proc. IEEE Int. Conf. Comput. Vis. **10**, 3037–3046 (2017)
12. Goodfellow, I.J., Pouget-Abadie, J., Mirza, M.: Generative adversarial networks. Adv. Neural. Inf. Process. Syst. **3**, 2672–2680 (2014)
13. Koch, G., Zemel, R., Salakhutdinov, R.: Siamese neural networks for one-shot image recognition. ICML Deep Learning Workshop 2 (2015)
14. Snell, J., Swersky, K., Zemel, R.S.: Prototypical networks for few-shot learning. In: Advances in Neural Information Processing Systems, vol. 6 (2017)
15. Oriol, V., Blundell, C., Lillicrap, T.: Matching networks for one shot learning. In: Advances in Neural Information Processing Systems, vol. 12, pp. 3630–3638 (2016)
16. Finn, C., Abbeel, P., Levine, S.: Model-agnostic meta-learning for fast adaptation of deep networks. In: International Conference on Machine Learning, vol. 3, pp. 1126–1135 (2017)
17. Ravi, S., Larochelle, H.: Optimization as a model for few-shot learning. In: International Conference on Learning Representations (2017)