https://doi.org/10.48047/AFJBS.6.12.2024.450-463

African Journal of Biological Sciences
Journal homepage: http://www.afjbs.com
Research Paper Open Access

PROVISIONING A COMPREHENSIVE SURVEY ON CERVICAL CANCER WITH FEW-SHOT LEARNING MODEL

Venkata Anupama Chitturi [1] **Dr Dharmaiah Devarapalli [2]**
2212031108@kluniversity.in drdharmaiah@kluniversity.in
[1], [2]Department of Computer Science and Engineering,
Koneru Lakshmaiah Education Foundation, Vijayawada-520002 A.P, India.

Article History

Volume 6 Issue 12, 2024
Received: 25 May 2024
Accepted : 25 June 2024
doi:
10.48047/AFJBS.6.12.2024.450-463

Abstract-Cervical cancer is a common malignancy encountered among many people in recent days. Early identification and detection can significantly aid in the management and treatment that follows. As advances in artificial intelligence (AI) and deep learning (DL) approaches have been made, an expanding range of deep learning-based (computer-aided detection (CAD) techniques have been used for cervical cytology screening. Firstly, we present a brief overview of the biological and medical information related to cervical cytology, as a thorough understanding of biomedicine can significantly impact the advancement of computer-aided detection (CAD) systems. Next, an extensive analysis of few-shot learning analyzed by many researchers is discussed. Furthermore, a synopsis of image analysis methods and uses is offered, such as cervical entire slide image diagnostic, cell region segmentation, detection of aberrant cells or areas, and identification of the cervical cell. In conclusion, they address current challenges and prospective avenues for further investigation into automated cervical cytology screening research.
Keywords-cervical cancer, few-shot learning, deep learning, prediction, cytology

1. Introduction

One common illness that seriously endangers the health of women is cervical cancer. In terms of both incidence and death, it is the fourth most frequent type of cancer [1]. Over 340,000 persons worldwide lost their lives to Cervical cancer in 2023 accounted for about 600,000 new cases of the illness that were detected [2]. In different countries andregions, the incidence and mortality of cervical cancer can be influenced by various factors such as the quality of healthcare provided, the implementation of screening and preventive measures, local lifestyles, and environmental conditions [3] – [5]. One type of malignant tumour that develops from the cervix is cervical cancer, which puts women's lives and health at risk. Adenocarcinoma and squamous cell carcinoma (SCC) are the two primary forms of cervical cancer [6]. SCC accounts for about 90% of instances of cervical cancer [7]; most of these cases originate from cells found in the cervix's external layer and begin in the transformation zone [8]. The most prevalent disease linked to the human papillomavirus (HPV) is cervical cancer, of which over 95% are the result of ongoing infection with certain HPV strains [9]. High-risk HPV refers to at least 13 known strains of HPV, the most prevalent of which are HPV 16 and 18. These strains can continue and develop into cancer [10].

As seen in Fig. 1, cervical cancer has a protracted precancerous stage and a continuous course of progression. Modifications to the nucleus are the main characteristic of precancerous cells [11]. For instance, the nuclear-to-cytoplasmic ratio (N/C ratio) rises with nuclear expansion. One frequently observes both bi-nucleation and multi-nucleation [12]. Additionally, nucleoli are either rarely seen or barely noticeable when they are. The nuclear membrane has a highly uneven shape [13]. Early-stage cervical cancers may not show any symptoms at all, but when the illness worsens, symptoms such as pelvic pain, vaginal discharge, and irregular vaginal bleeding [14]. For cervical cancer therapy and prognosis, early identification is therefore crucial [15].

Preventive and effective management of cervical cancer through regular screening programs can make it one of the most treatable and preventable malignancies, according to solid evidence. Cervical cytology screening, three screening methods for cervical cancer have been recommended by the World Health Organization (WHO) [16]: HPV testing for high-risk HPV strains, visual examination with acetic acid (VIA), and HPV testing.Since cytological characteristics are essential indicators of cervical cancer, cervical cytology screening is now the norm everywhere in the world. The cervical cytology screening program has long required the painstaking, laborious, and error-prone process of manually identifying aberrant cells under a microscope [17]. In this regard, an increasing amount of automated screening tools have been put out to lessen the workload of cytopathologists and increase the effectiveness of their diagnoses. Because machine learning (ML) technology produces high-performance outcomes, it has been widely used to assess cytological images during cervical cytology screening. ML technology has advanced with digital image processing and artificial intelligence (AI) [18].

However, the advancement of human-machine collaboration is limited by the intricate image preparation and feature selection processes of classical machine learning systems. There has been a massive expansion in recent years because a machine learning subfield known as deep learning (DL) has been used in computer vision [19]. Deep learning automates the process of learning features from start to finish, eliminating the need for human feature design and selection. Analyzing medical images is only one area in which deep learning (DL) has made significant strides. Imaging modalities include thoracic, neural, cardiovascular, abdominal, and microscopic imaging are examples of medical imaging tasks that have all benefited from the successful application of DL solutions [20] – [21]. Additionally, the automatic picture processing for cervical cytology screening has been significantly sped by the advancement of deep learning.

To address the problems above, a thorough review of pertinent literature on cervical cytology was performed automatically. This study also offers background information on cervical cytology,comprising specific cell types [22], a synopsis of cervical cancer, and conventional cervical cytology screening protocols (TBS). The fact that a comprehensive comparison of various reporting terms is provided is interesting.This frequently leads to misunderstandings and hinders the creation of accurate and sensible DL models [23]. Furthermore, a detailed introduction to the particular duties involved in cervical cytology screening has been provided, in addition to the progressive advancement of automated screening techniques. Additionally, this survey has assembled the most extensive collection of publicly accessible cervical cytology imaging files. Additionally, this survey provides an overview of the most recent techniques for automated cervical cytology screening that use few-shot learning-based analysis, detection, and classification based on DL. This paper concludes with several challenges and opportunities that may lead to interest in the Internet of medical things; stain normalization, super-resolution images, integrating medical domain knowledge, and annotation-efficient learning are some of the routes that future research in cervical cytology screening may go [24].

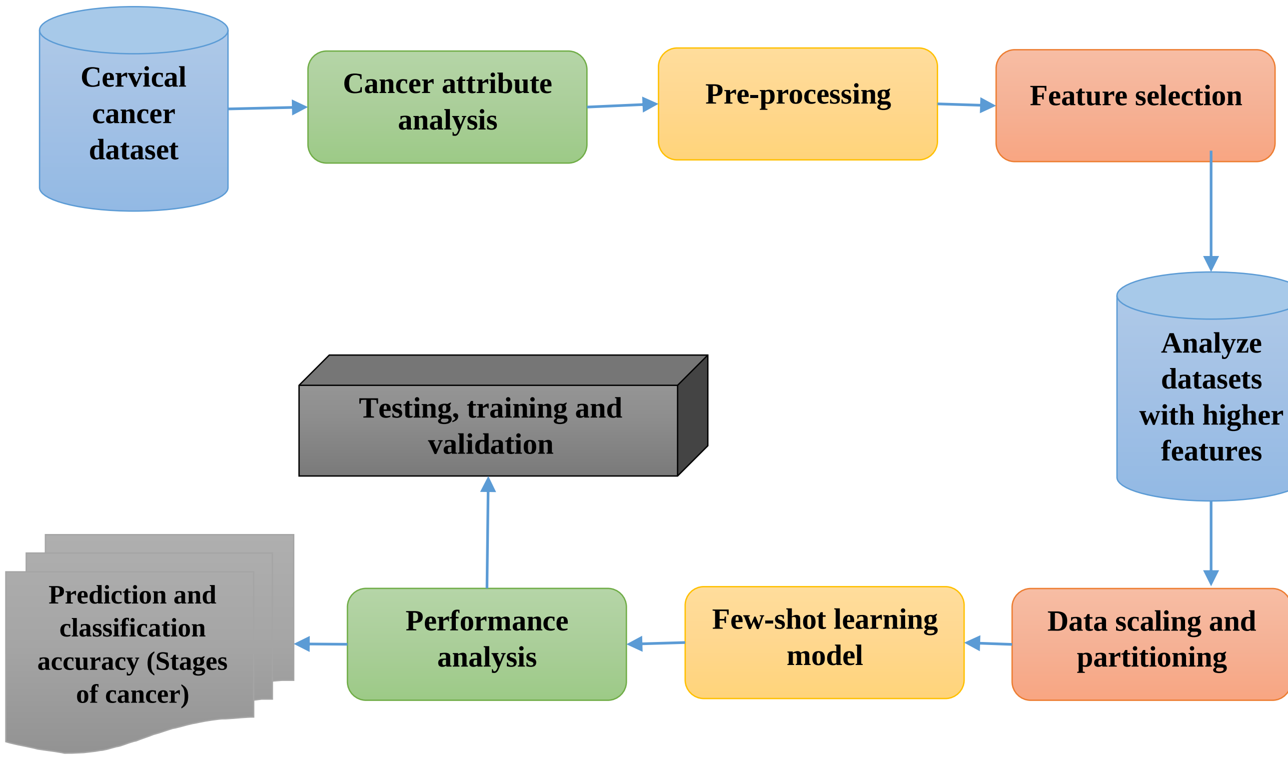

Fig 1 Cervical cancer detection

2. Few shot learning

In this section, they present a solution based on meta-learning for long-tailed datasets and sketch a few-shot picture classification problem. Here, $D_k = \{(x,y)_j\}_{j=1}^{n}$ defines each medical dataset when the notation $(x,y)j$ is used to denote a pair of images and the label that corresponds to them (ground truth). To represent the total of n medical datasets, let $D = \{D_1, D_2 \dots, D_n\}_{k=1}^{n}$. Each dataset is split into two groups: The photographs of classes with fewer representative photos (rare diseases) are included in the meta-test set ($D_{meta-test}$). Moreover, the remaining classes are selected to form the meta-train set ($D_{meta-train}$).The key takeaway from this is to improve your initial weights and make adjustments for problems when there is insufficient data by using relevant data that is abundantly available ($D_{meta-train}$). To tackle the data scarcity issue, researchers have formulated it as a few-shot learning issue.The ultimate objective is to train a model called $Y = f(x;\varphi)$, where φ represents the neural network parameters, to perform well on few-shot tasks taken from the $D_{meta-test}$ sample.Because for deep learning algorithms to generalize successfully, enormous volumes of data must be dispersed nearly evenly across all classes. However, they must achieve high accuracy when faced with a few-shot learning problem [25]. Thus, our approach aims to acquire more robust initial weights using the meta-train set to achieve increased accuracy, subsequently refined on unseen datasets.

2.1. Few-shot learning

The variables "k" and "n" in a k-way, n-shot learning problem stand for the total number of images and the number of classes an input needs to be classified. That can be learned for each class. For instance, the model is given 10 photos in a two-way, five-shot learning task; there are five for every class,letting it pick up new skills and classify the test photos into two groups. The learning process, as it is divided into multiple episodes, is described by the meta-learning paradigm.Each of these includes the phases for meta-training and meta-updating. The primary learner is represented by the model $y = f(x; \varphi)$; in this case, x represents the input, and φ denotes the weight of network a. Because of its effectiveness and simplicity, we utilize the Reptile method to train the primary learner. During meta-training, the classification model$y = f(x; \varphi_k)$, where $\varphi_k = \varphi$ is initialized for each job k in m such tasks using the meta parameter $\varphi^{'}$.

$$\varphi^{'} \leftarrow \varphi + \in \frac{1}{m}\sum_{k=1}^{m}(\varphi_k^{'} - \varphi) \qquad (1)$$

Following the modelling of every job using the meta-train set ($D_{meta-train}$), Adam is used to update the task parameter φ_k for each task. Ultimately, the learned task parameters φ_k is used to meta-update the meta parameter φ, yielding φ and an update rule. Assume that the test and training splits of the $D_{meta-test}$t are represented by test($D_{meta-test}$) and train($D_{meta-test}$). The total number of n_k photographs (k images per class) will be utilized to fine-tune for each n −way, k −shot task sampled from the train ($D_{meta-test}$), the weights derived in Eq. [2] for a decreasing number of iterations (h) using Adam with learning rate (α), resulting in φ. The correctness of the model is then assessed using φ on the test ($D_{meta-test}$), which will comprise 'n' images. We updated the model's weights using cross-entropy loss during the meta-training and meta-testing phases. The cross-entropy loss for a given job T_l can be found mathematically by:

The pair of labels and inputs (x_i, y_i) covers the augmentation phases utilized; Algorithm 1 addresses the step of fine-tuning and meta-training.

$$L_{T_l}(f_\emptyset = -\sum_{x_i y_i \sim T_l} y_i \log\left(f_\emptyset(x_i)\right) + (1 - y_i)\log(1 - f_\emptyset(x_i)) \qquad (2)$$

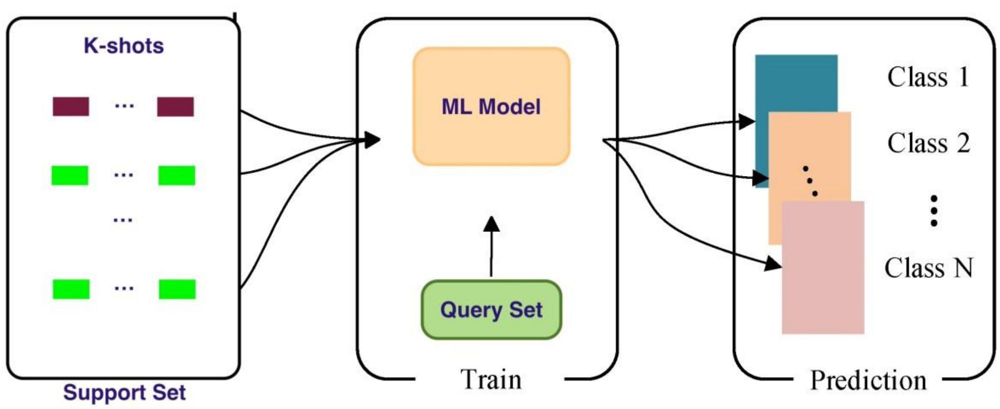

Fig 2: k-shots in Meta-data learning

A similar approach was put forth in the study [26], concentrating on challenging tasks to enhance model performance and exclusively applying the framework to the skin lesion categorization problem. As stated in the study [26], the technique, which Difficulty Aware Meta-Learning (DAML) employed, requires computing second-order gradients and does not converge quickly. The MAML states that the model uses the transductive mode instead of considering second-order gradients. Some existing studies adopt a new strategy by leveraging sophisticated augmentation methods to improve the model's functionality.

Algorithm 1:k-few shot Algorithm

Input: Training samples, learning rate, distribution parameters, adaptation steps, few-shot tasks, iterations
Output:accuracy based on a few shots
1. Initialize network weight randomly;
2. **for** iteration $i = 1, \dots, n$**do**
3. Train samples randomly from the training set;
4. **for** iteration $i = 1, \dots, n$**do**
5. **if** pre-processing, then //augmentation
6. Validate alpha and meta-training data samples;
7. **end**
8. Execute training parameters;
9. Update learning rate and loss function;
10. **end**
11. Perform value updation with φ';
12. **end**
//Testing
13. **for** iteration $i = 1, \dots, n$**do**
14. Perform k-few shot task;
15. Partition samples based on training and testing;
16. Load network weight using φ';
17. **for** all iteration $i = 1, \dots, n$ $\boldsymbol{do}$
18. update φ' using loss function;
19. Perform optimization; //adam
20. **end**
21. compute training and testing accuracy;
22. **end**

It demonstrates how easily trainable and versatile it is. They apply this methodology to any medical dataset. One more study that focuses just on the classification of skin lesions is called Meta-Derm Diagnosis [26]. It builds on the work of DAML utilizing Reptile and increases accuracy even more using Group Equivariant Convolution. The framework cannot be used with other medical datasets that lack symmetry since meta-derm diagnosis uses Group Equivariant

Convolution. We do not rely on any such data-specific characteristics in our framework. Our approach outperforms the ISIC 2018 skin lesion dataset results compared to Meta-DermDiagnosis without Group Equi-variant Convolution. Since the Group Equi-variant Convolution method uses characteristics unique to the data to increase accuracy, we disregard the comparison [27].

3. Evaluation process

This section presents a general approach for evaluating the performance of a classifier on the N −way K −shot image classification tasks. Numerous episodes make up the entire evaluation procedure. Using K samples per class from the novel label space, we first select N classes at random to create. In each episode, a support set (D_S)and M examples from the remaining samples of those N classes are used to create a query set (D_Q) [27]. The last classifier in DQ is utilized to forecast sample labels that can be made using the primary dataset and support set. The symbol $acc(e)$symbolizes the episode's categorization accuracy on earth, and the average classification accuracy over all episodes can be used to gauge how well a learning system is performing.

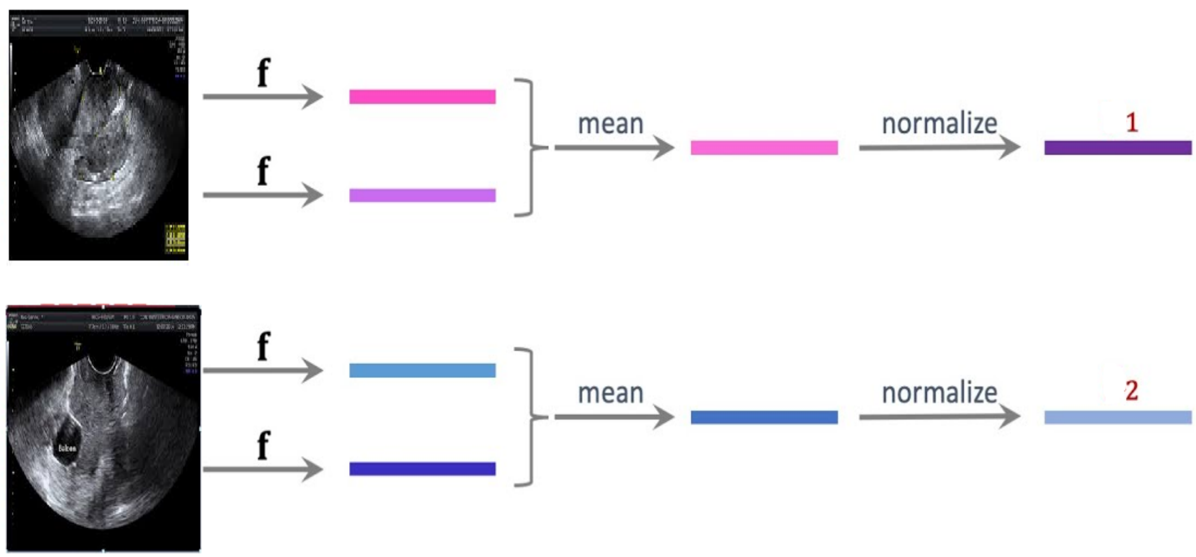

Fig 3 Cervical cancer prediction with few-shot learning

3.1. Meta-learning for few-shot learning

Enhancing the performance of models and intense on novel problems with limited sample sizes is the aim of meta-learning for few-shot image categorization. Several meta-learning strategies have been put forth in light of few-shot learning's explosive growth. This is a thorough synopsis of the most recent meta-learning research and its developments. To facilitate comprehension for novices, researchers adhere to the prevailing approach and divide meta-learning techniques into three categories: model-, optimization-, and metric-based [28]. In addition, they provide additional few-shot learning techniques for comparison.

3.2. Metrics based few shots Meta-learning

The goal of metric-based meta-learning techniques is to derive a distance metric that can quantify sample similarity with high accuracy and efficiency while guaranteeing its applicability to novel learning tasks.The acquired measure for single-shot image categorization tasks should adhere to the principles that allow instances from the same class or a different one to be separated by a short (or considerable) distance. The Siamese network is one of the most popular metric-based techniques for one-shot picture classification. Initially, signature verification was mentioned when the term "Siamese" was introduced [29], and the Siamese network's main structure was first presented with the challenge of estimating fingerprint similarity. The author used a pair of identical VGG-styled convolutional layers with shared weights to extract high-level features from two input pictures by calculating the weighted L1 distance between the two feature vectors. A score generated by the network finally indicates the likelihood that the two pictures are from the same class.A Siamese network with an attention component was proposed by [29] that performs the similarity calculation between two feature vectors using an attention kernel function. A multi-resolution Siamese network using a hybrid training strategy that mixes multiple kernel size streams into a single layer was proposed to close the gap between one-shot photo identification and regular classification.

A matching network [30] is another potent metric-based meta-learning technique that encodes support and query images using various networks.An LSTM with an attention kernel is used for query image embedding. A bidirectional long-short-term memory (LSTM) is utilized to facilitate the dependence on the support set DS for support image embedding.The cosine similarity between the support and query images is calculated using the attention kernel, which is then normalized using a softmax function. Once the attention kernel has weighted the one-hot encoded support picture labels, the total of those labels is the output of the matching network. An attentive matching network (AMN) was developed. A feature-level attention method was introduced to prioritize characteristics that can more accurately reflect variations across classes and an additional cosine loss function for optimization.
A distance metric called the earth mover's distance (EMD) is used to assess the similarity in [31] DeepEMD technique to obtain ideal matching picture regions. A cross-reference method is introduced to generate the component weights for the EMD formulation. Next, the network's EMD layer is integrated for end-to-end training. This led to the proposal of the deep Brownian distance covariance (DeepBDC) technique, which employs few-shot learning with the BDC measure. Mixture-based feature space learning, or MixtFSL, is the proposed method for learning discriminative feature representations. This approach entails online learning of blending models and feature representations. As opposed to the few-shot classification techniques, which take one feature vector per image,

3.3. Model-based Meta-learning

Model-based approaches primarily concentrate on model designs to facilitate quick learning, modifying model parameters according to given tasks. Convolutional neural networks (CNNs), recurrent neural networks (RNNs), and long short-term memory (LSTM) are a few architectures that are widely employed in model-based approaches [32]. Depending on the types of model architectures,memory-based, quickly adaptable, and using different models are further subdivided under these model-based techniques. The memory-augmented neural network (MANN) was introduced. A well-known memory-based technique that improves using the neural Turing machine (NTM) to adjust tasks.A neural network called an NTM incorporates a component of external memory into its educational process to access and recover information that has already been stored. NTM is comprised of a controller that uses multiple read and write heads to interface with an external memory module. The NTM scheme is displayed. A novel addressing method for MANN is suggested, called least recently used access (LRUA); it stores memories in the designated memory location utilized the least or the most recently. MANN can access them for subsequent categorization using the associated information about representation-class labels stored in an external memory. A memory-augmented matching network (MAMN) that combines MANN and a matching network was proposed.

By summing the distances of class-wise samples, MAMN generates weighted class prototypes, which lessen the bias on class prototypes brought on by the skew in the data distribution. Memory matching network (MM-Net) [33] is an additional memory-based meta-learning technique created by combining memory modules from the matching network of the key-value memory network. Every support set is encoded into memory slots by MM-Net, generalizing it in contrast to traditional one-shot learning approaches, enabling the generation of a single model independent of the number of shots and categories. An architecture for universal model-based meta-learning called the Simple Neural Attentive Learner (SNAIL) [34] combines a soft attention mechanism with temporal convolution.Access to specific bits of information is made possible by the soft attention and high-bandwidth memory access provided by the temporal convolution. Thanks to this combination, models can better use information from prior experiences.Task learners and meta-learners are the two components of conditional neural processes (CNPs) proposed by [35] and are similar to SNAIL. The task learner processes the aggregated representations to provide predictions, while the meta-learner aggregates representations of the support set to create a memory value.

3.4. Optimization-based meta-learning

A significant area of study in the field of few-shot image categorization is optimization-based meta-learning techniques. This algorithm aims to improve the gradient descent direction or the initialization model by utilizing the meta-learning architecture. This is accomplished by using episodic training to optimize the starting settings, and only a limited set of training samples can be used in an optimization process. Optimization-based approaches typically comprise a meta-learner taught on task distributions and trained for a particular task, a task-specific learner.The first technique for learning an initialization is the model-agnostic meta-learning proposal by [36]. The primary concept of MAML is the application of gradient-based learning rules to allow the parameters of a model to adjust to novel, untested tasks swiftly. In the meta-training stage, MAML tries to change both the global initialization and the task-specific parameters iteratively. The MAML scheme is displayed. The primary advantage of MAML is its adaptability to several application areas, including addition classification, regression and reinforcement learning. An LSTM-based meta-learner was proposed by [37] to discover the proper beginning values for a classifier's precise task-specific optimization and the parameters of the task-specific learner.

This was done in reaction to the constraints of neural networks trained on few-shot learning tasks using gradient-based optimization. Low-shot learning using task-aware feature embeddings (TAFE-Net), a novel approach, focuses primarily on optimizing embedding task-specific features using a meta-learner's generic embedding. TAFE-Net is composed of a prediction network and a meta-learner. The meta-learner produces task-aware feature embedding in the prediction network's task-specific feature layers.A meta-transfer learner (MTL) approach was presented by [38]. It aims to combine transfer learning and meta-learning to produce task-specific feature extractors. Pre-trained feature embedding is subjected to shifting and scaling operations to freeze the feature extractor in MTL.Moreover, MTL follows the same fine-tuning procedures as earlier research. Additionally, this study described a unique, challenging task meta-batch approach that collected additional samples from the classes in which the classifier failed, focusing greater attention on heavy workloads. Latent embedding optimization (LEO), in a unique method

presented by [39], model parameters are learned as a low-dimensional latent representation, and meta-learning via optimization is done in this domain.
This algorithm addresses the challenges associated with optimization on large parameter spaces, like those that MAML must deal with [40]. Like MAML, LEO comprises two training loops: an outer loop that updates global common initializations and an inner loop that learns task-specific values. Samples are passed through a related network and an encoder to create an instantiated low-dimensional latent embedding of the parameters in the model. Secret codes are generated by the encoder from the support set. A probability distribution is obtained in a dimension below latent codes by concatenating these hidden codes pairwise and feeding the resulting data into a relation network. Ultimately, the decoder generates task-specific starting parameters that can be back-propagated for modification.

3.5. Other approaches

Using what has been learnt in one activity to improve learning in another is known as transfer learning. Using information from another network can be a good choice in the few-shot photo categorization case when there is not enough original data to train a deep neural network from the beginning.Transfer learning involves a far smaller learning experience than meta-learning. Using the transfer learning approach, learnt intrinsic representations were retrieved by [41] from the same type of objects across several domains to handle few-shot hyper-spectral image classification challenges.A novel few-shot transfer learning technique was proposed to classify synthetic aperture radar images. It transfers features from one network to another using a connection-free attention module. Regarding issues with few-shot fine-grained visual classification, the author introduced a two-phase learning technique called trans-transfer learning.

Table 1 Comparison of various existing approaches

Ref	**Methods**	**Dataset**	**Dataset classes**	**Outcomes**
Do et al., (2022)	Faster RetinaNet CNN	Cervical DCLL dataset with 27 973 images	Two classes	(Macro Average Precision) mAP = 0.26
Faradonbe et al., (2020)	Inception v3 + YOLOv3 + smoothing	12000 cervical images with ten different categories	Ten classes	mAP = 0.63
Gaiwad et al., (2022)	YOLOv4	Tian-Chi cervical dataset	Six classes	AP = 0.54, Precision = 0.71 Accuracy = 0.91
Li et al. (2021)	Improved BSNet	12000 cervical images with ten different categories	Ten classes	mAP = 0.65

Li et al. (2022b)	Enhanced YOLOv3	2000 private images	Four classes	mAP = 0.81
	Pre-trained VGG-16 with FCN	73 private images	Ten classes	mAP = 0.78
Gu et al. (2018)	Masked R-CNN	2700 normal images, 494 abnormal images and 85 high-graded images	Two classes	AP = 28 and sensitivity = 0.30
Qin et al. (2020)	R-FCN	65 cervical images	Four classes	mAP = 0.57, accuracy = 0.91 and sensitivity = 0.91
Chen et al. (2021)	Cascaded dense R-CNN	Tian-Chi cervical dataset	Two classes	(Average Precision) AP = 0.93
Cai et al. (2018)	HSDet	Microscopic images with 49000 samples	Two classes	AP = 0.97
Afrasiyabi et al., (2021)	Faster R-CNN with a few-shot learning model	2000 private images	Eleven classes	AP = 0.50
Bian et al. (2021)	FPN + Faster R-CNN and transfer learning	Herlev dataset	Ten classes	mAP = 0.50
Brauwers et al., (2023)	Attention Mechanism and RetinaNet	Microscopical images with 49000images	Six classes	mAP = 0.61
Lalonde et al., (2020)	Alignment + Few shot + R-CNN	Private images	Four classes	mAP = 0.25

When there is no relationship between the source and destination domains, knowledge transfer may occasionally fail and even result in negative transfer. Using the analogy technique to control negative transfer properly is a crucial component of Analogical Transfer Learning (ATL) and was developed to solve this issue. The underlying problem with few-shot image classification is that a small number of training instances can lead to models needing more fit. Several researchers have developedseveral techniques for data augmentation to increase the sample size and avoid over-fitting during training. The popular Generative Adversarial Nets (GAN) were first presented by [42]. They comprise a discriminator for discriminating and a generator for producing similar images. The author suggested producing samples based on GAN for specific tasks to improve the suitability of these generated samples for few-shot learning.

4. Challenges

Although meta-learning techniques have demonstrated encouraging results in few-shot picture categorization, some significant issues still need to be resolved. These are the current problems and recommended avenues for future research to be explored [43]—the accessibility of data and the complexity of calculation.A thousand or more categories are usually included in a large dataset in picture classification. Large amounts of data and processing power are also needed for meta-learning systems, although gathering enough data in few-shot scenarios can be pretty tricky. We could require hundreds of massive datasets for meta-learning deep testing! Process-wise, this could be pretty challenging and slow.

Model selection: Since no one-size-fits-all solution exists, choosing the suitable model is crucial. Model selection is essential in few-shot image classification scenarios since the model's training set frequently over-fits. The model may function effectively with the basis set, but more must be applied to other tasks [44].

Portability: Meta-learning models can apply knowledge gained to other activities. The degree to which the jobs are comparable determines whether transferability is successful. It can be challenging to successfully transfer gained information to new tasks when they differ significantly from old ones, as in the case of cross-domain activities [45].

Task dependency: A particular collection of activities or domains is the target audience for most meta-learning systems. They won't function well on novel assignments or domains that differ significantly from training ones. Enhancing the capacity for generalization in meta-learning might be challenging [46].

Interpretability: Interpretability, or the capacity to comprehend a model's operation, is a crucial component of neural techniques. Regretfully, all neural methods have the potential to be very difficult to interpret, making it difficult to comprehend how the system learns to recognize and make predictions or judgments. It may be difficult to debug, diagnose, and enhance the performance of models because of this problem [47].

4.1. Future research directions

Improving learning of generalized features to overcome the primary difficulty with few-shot learning, Meta-learning uses shared data from activities that have already been finished to help with tasks that haven't been observed yet in contrast to a small sample size learning approach [48]. Nonetheless, most current meta-learning techniques aim to teach researchers discriminative characteristics through attention mechanisms, multitasking learning, data augmentation, and more.Two critical areas of research are creating creative ways to learn characteristics that are more applicable to various fields and using assessment tools to choose and evaluate the features that have been realized [49].

Increasing steadiness: While meta-learning in few-shot picture categorization continues to advance, there still needs to be a problem with specific meta-learning techniques achieving cutting-edge results on specific datasets but underperforming on other benchmarks [50]. For instance, global class representation (GCR), a metric-based meta-learning technique, performed superbly on Omniglot, but on miniImageNet, it could not match other non-metric-based approaches. Looking at stable models further will be very helpful.

5. Research objectives

The primary objective of upcoming research is to adopt few-shot learning for modelling prediction algorithms and to learn the raw and generalized data towards a new instance (unseen instances). It generally uses prior leveraging knowledge from various associates' tasks to generalize new tasks effectually. The proposed model must address the above-given research challenges and provide optimal solutions.

6. Conclusion

This study surveys contemporary few-shot learning and meta-learning studies related to visual interpretation. It begins by outlining the broad few-shot learning approaches based on the research literature, and it then moves on to discuss one of the critical strategies, meta-learning. The primary categories of contemporary meta-learning approaches are distinguished by researchers: optimization-based, model-based, and metric-based. It showcases the most advanced techniques in each category and conventional and innovative approaches. Additionally, it showcases the most advanced results achieved by the literature methodologies on reputable datasets. Our research leads us to conclude that meta-learning has certain restrictions, difficulties, and shortcomings. Additionally, researchers offer future directions for meta-learning from application, efficacy, and generalization perspectives.

REFERENCES

[1] Afrasiyabi A, Lalonde J, Gagné C (2021) Mixture-based feature space learning for few-shot image classification. In: ICCV, pp 9021–9031

[2] Afrasiyabi A, Larochelle H, Lalonde J et al. (2022) Matching feature sets for few-shot image classification. In: CVPR, pp 9004– 9014

[3] Afrasiyabi A, Lalonde J, Gagné C (2020) Associative alignment for few-shot image classification. In: ECCV, pp 18–35

[4] Alzubaidi L, Zhang J, Humaidi AJ et al. (2021) Review of deep learning: concepts, CNN architectures, challenges, applications, future directions. J Big Data 8(1):53

[5] Baik S, Choi J, Kim H et al. (2021) Meta-learning with task-adaptive loss function for few-shot learning. In: ICCV, pp 9445– 9454

[6] Bian W, Chen Y, Ye X et al. (2021) An optimization-based meta-learning model for MRI reconstruction with the diverse dataset. J Imaging 7(11):231

[7] Brauwers G, Frasincar F (2023) A general survey on attention mechanisms in deep learning. IEEE Trans Knowl Data Eng 35(4):3279–3298

[8] Cai Q, Pan Y, Yao T et al. (2018) Memory matching networks for one-shot image recognition. In: CVPR, pp 4080–4088

[9] Chao X, Zhang L (2021) Few-shot imbalanced classification based on data augmentation. Multimed Syst, pp 1–9

[10] Chen Y, Liu Z, Xu H et al. (2021) Meta-baseline: exploring simple meta-learning for few-shot learning. In: ICCV, pp 9042–9051

[11] Ding G, Han X, Wang S et al. (2022a) Attribute group editing for reliable few-shot image generation. In: CVPR, pp 11,184–11,193

[12] Dong J, Wang Y, Lai J et al. (2022) Improving adversarially robust few-shot image classification with generalizable representations. In: CVPR, pp 9015–9024

[13] Do J, Yoo M, Kim S (2022) A semi-supervised sar image classification with data augmentation and pseudo labelling. In: ICCE-Asia, pp 1–4

[14] Do J, Yoo M, Kim S (2022) A semi-supervised sar image classification with data augmentation and pseudo labelling. In: ICCE-Asia, pp 1–4

[15] Fallah A, Mokhtari A, Ozdaglar A (2020) Provably convergent policy gradient methods for model-agnostic meta-reinforcement learning. arXiv preprint arXiv:2002.05135

[16] Faradonbe SM, Safi-Esfahani F, Karimian-kelishadrokhi M (2020) A neural turning machine (NTM) review. SN Comput Sci 1(6):333

[17] Finn C, Abbeel P, Levine S (2017) Model-agnostic meta-learning for fast adaptation of deep networks. In: Precup D, Teh YW (eds) ICML, pp 1126–1135

[18] Gaikwad M, Doke A (2022) Survey on meta-learning algorithms for few-shot learning. In: ICICCS, pp 1876–1879

[19] Gao K, Sener O (2020) Modeling and optimization trade-off in meta-learning. In: NeurIPS, pp 11,154–11,165
[20] Gidaris S, Komodakis N (2019) Generating classification weights with GNN denoising autoencoders for few-shot learning. In: CVPR, pp 21–30
[21] Gu J, Wang Z, Kuen J et al. (2018) Recent advances in convolutional neural networks. Pattern Recognit 77:354–377
[22] Guo Y, Codella N, Karlinsky L et al. (2020) A broader study of cross-domain few-shot learning. In: ECCV, pp 124–141
[23] Han G, Ma J, Huang S et al. (2022) Multimodal few-shot object detection with meta-learning based cross-modal prompting. arXiv preprint arXiv:2204.07841
[24] Hu T, Tang T, Lin R et al. (2020) A simple data augmentation algorithm and a self-adaptive convolutional architecture for few-shot fault diagnosis under different working conditions. Measurement 156(107):539
[25] Huang H, Zhang J, Zhang J et al. (2019) Compare more nuanced: pairwise alignment bilinear network for few-shot fine-grained learning. In: ICME, pp 91–96
[26] Kang B, Liu Z, Wang X et al. (2019) Few-shot object detection via feature reweighting. In: ICCV, pp 8419–8428
[27] Karunaratne G, Schmuck M, Le Gallo M et al. (2021) Robust high-dimensional memory-augmented neural networks. Nat Commun 12(1):2468
[28] Köksal A, Schick T, Schütze H (2022) Meal: stable and active learning for few-shot prompting. arXiv preprint arXiv:2211.08358
[29] Li X, Yu L, Fu C et al. (2020) Revisiting metric learning for few-shot image classification. Neurocomputing 406:49–58
[30] Li X, Wu J, Sun Z et al. (2021) Bsnet: Bi-similarity network for few-shot fine-grained image classification. IEEE Trans Image Process 30:1318–1331
[31] Li X, Yang X, Ma Z et al. (2023) Deep metric learning for few-shot image classification: a review of recent developments. Pattern Recognit 138(109):381
[32] Li W, Liu X, Bilen H (2022b) Cross-domain few-shot learning with task-specific adapters. In: CVPR, pp 7151–7160
[33] Liu Y, Zhang H, Zhang W et al. (2022) Few-shot image classification: current status and research trends. Electronics 11(11):1752
[34] Liu W, Chang X, Yan Y et al. (2018) Few-shot text and image classification via analogical transfer learning. ACM 9(6):71:1– 71:20
[35] Li W, Xu J, Huo J et al. (2019c) Distribution consistency based covariance metric networks for few-shot learning. In: AAAI Conference on Artificial Intelligence, pp 8642
[36] Mai S, Hu H, Xu J (2019) Attentive matching network for fewshot learning. Comput Vis Image Underst 187(102):781
[37] Mangla P, Singh M, Sinha A et al. (2020) Charting the right manifold: manifold mixup for few-shot learning. In: WACV, pp 2207–2216
[38] Moon J, Le NA, Minaya NH et al. (2020) Multimodal few-shot learning for gait recognition. Appl Sci 10(21):7619
[39] Nguyen VN, Løkse S, Wickstrøm K et al. (2020) SEN: a novel feature normalization dissimilarity measure for prototypical few-shot learning networks. In: ECCV, pp 118–134
[40] Osahor U, Nasrabadi NM (2022) Ortho-shot: low displacement rank regularization with data augmentation for few-shot learning. In: WACV, pp 2040–2049

[41] Parnami A, Lee M (2022) Learning from few examples: a summary of approaches to few-shot learning. arXiv preprint arXiv:2203.04291
[42] Peng Z, Li Z, Zhang J et al. (2019) Few-shot image recognition with knowledge transfer. In: ICCV, pp 441–449
[43] Qiao S, Liu C, Shen W et al. (2018) Few-shot image recognition by predicting parameters from activations. In: CVPR, pp 7229–7238
[44] Qin T, Li W, Shi Y et al. (2020) Diversity helps: unsupervised few-shot learning via distribution shift-based data augmentation. arXiv preprint arXiv:2004.05805
[45] ReifM, Shafait F, Dengel A (2012)Meta-learning for evolutionary parameter optimization of classifiers. Mach Learn 87(3):357–380
[46] Satrya WF, Yun J (2023) Combining model-agnostic meta-learning and transfer learning for regression. Sensors 23(2):5
[47] Shi G, Chen J, Zhang W et al. (2021) Overcoming catastrophic forgetting in incremental few-shot learning by finding flat minima. In: NeurIPS, pp 6747–6761
[48] Singh R, Bharti V, Purohit V et al. (2021) Metamed: few-shot medical image classification using gradient-based meta-learning. Pattern Recognit 120(108):111
[49] Sun N, Yang P (2023) T2L: trans-transfer learning for few-shot fine-grained visual categorization with extended adaptation. Knowl Based Syst 264(110):329
[50] Tian Y, Wang Y, Krishnan D et al. (2020) Rethinking few-shot image classification: A good embedding is all you need? In: ECCV, pp 266–282