https://doi.org/10.48047/AFJBS.6.12.2024.2170-2178

AFJBS African Journal of Biological Sciences
Journal homepage: http://www.afjbs.com

Research Paper Open Access

A deep learning-based analytical survey on Few Shot Image Dataset Cleaner

Vidya Mali [1,]
Sharad Institute Of Technology college of Engineering, Yadrav , 416115,vidyamali2023@gmail.com
Dr. Shashidhar Gurav[2]
Sharad Institute Of Technology college of Engineering, Yadrav , 416115,gurav.shashidhar@gmail.com
Dr. Amit Chinchawade[3] **,**
Sharad Institute Of Technology college of Engineering, Yadrav , 416115,amitchinchwade@sitcoe.org.in
Varsha Jujare[4] **,**
Sharad Institute Of Technology college of Engineering, Yadrav , 416115,varsha.jujare@gmail.com

Article History
Volume 6, Issue 12, 2024
Received: 30 May 2024
Accepted : 30 June 2024
Doi:
10.48047/AFJBS.6.12.2024.2170-2178

Abstract.
AI-based image analysis and disease identification are getting more and more popular. Examples of these disorders include mouth cancer, cervical cancer, retinal glucose analysis, etc. Capturing images of the affected areas with customized camera modules is a common method of gathering data. Like any other data source, a technique that is prone to error and may contain undesirable objects and regions that need to be cleaned up by removal is known as a removal procedure. Outliers in these kinds of datasets can have a detrimental effect on the evaluation of the effectiveness of machine learning models. It would take a lot of effort to manually clean data, especially if it was gathered from multiple sources. As a result, cleaning the data is essential before model training. In this study, we incorporate reviewing earlier works to unleash the gaps and ideas behind the few-shot image classification models. This research paper mainly focuses on the identification of the possibility of implementing the hybrid model to clean the few-shot images. This hybrid model is built using two neural network models, for example, transformers. Generators or by using neural networks.

Keywords: Few-shot images, Adaptive Generators, Deep learning models, Transformers.

1 Introduction

Lately, there has been a swift advancement in remote sensing technology, including an increased variety of platforms and data-collecting methods. Numerous sectors, including land use, vegetation cover, pest monitoring, and urban planning, depend on the data analysis and processing from remote sensing. The relationship between information communication technology and remote sensing has led to many advances in the field of target recognition, object tracking as remote sensing images, and the possibility of spatial data reconstruction. These advances can be made using remote sensing images. Deep learning is a vital component of intelligent remote sensing applications as the primary technical support.

As is well known, deep learning is a common model example of a data-hungry learning technique that needs a large number of annotated images to handle computer vision tasks. Even other auxiliary learning techniques like self-supervised, semi-supervised, and unsupervised learning have become more popular, Supervised learning, which is based on data that has been labeled, is still the implementation strategy that is utilized the most frequently in real-world scenarios. The data for images obtained from remote sensing scenarios are often high resolution, and gathering, transmitting, and annotating all the samples on a wide scale is challenging and expensive. Therefore, it becomes sense to take data quality assessment into account to carry out effective learning based on sparse but extremely informative remote sensing data.

Few-shot learning, an addition to the present deep learning paradigm, is sometimes referred to as learning with limited data. Metric learning, transfer learning, and initial parameter optimization are three strategies used in few-shot learning techniques to assess the generalized learning ability of neural networks using weakly labeled data. Determining the similarity of image feature vectors—which is typically based on the Euclidean, cosine, Mahalanobis, etc. distances—is the main goal of the metric learning approach. Based on this, the classification and recognition steps are completed using nearest neighbor comparison. The transfer learning approach essentially requires at least two domains, with the source domain having a sufficient amount of annotated data, in order to train and obtain strong representation ability. To make appropriate modifications to the specifications, the network shall be transferred to the new domain. To begin the process of optimizing parameters, it is necessary to construct several different tasks and then employ meta-tasking to ascertain the optimal network settings at the time of task update. In several areas and industries, such as industry, healthcare, and agriculture, a few-shot studies and applications have lately surfaced. These works mostly use random sampling of few-shot data, ignoring the impact of data quality and relying only on numerous repeated experiments to counteract this significant problem. On the other hand, the fact of the matter is that a small amount of data that is either redundant or of poor quality can never yield a practical solution. In light of this, the potential of few-shot learning ought to be concentrated on researching the information contribution of samples to accomplish the goal of data-efficient learning.

Regarding the evaluation of picture data quality, there are essentially two groups. In the first case, there is a high-level task-oriented assessment of the picture information quality, which includes recognition and detection. In the second case, there is a visual perception-oriented assessment of the image quality, which includes distortion evaluation. The goal is to receive a high score on the visual perception-oriented picture quality evaluation, which assesses whether or not noise or distortion is present in the picture and how well it matches human perception. The intent and human perception are comparable, and this assessment's score reflects that. In the final analysis, there are a large number of works on distortion evaluation that fall into one of three categories: complete, reduced, or insufficient reference. The no-reference approach does not require a reference image in the situation where the full and reduced reference addresses use the difference between the reference and distorted image to calculate the visual quality score of the distorted image. However, these researches are not focused on solving the real-world issues in remote sensing applications. An assessment of the picture quality relies on visual perception and a recent task on feature representation of image data and information contribution has been done. The aim is to analyze data redundancy and information difference using information entropy for high-level recognition in picture quality assessment. These studies are crucial for exploring data-efficient learning strategies in real-world scenarios.

[1] A hybrid few shot divide and glow (FSDG) Net was proposed by Rizwan Khan, Qiong Liu, et al. to improve low- and images with inadequate lighting, both with and without the

application of paired training data supervision. A sizable training dataset is necessary for the current deep learning-based techniques, and some of them can only compete when paired training datasets are used as a guide. The authors separate the input images into illumination and reflection for the proposed FSDG method and provide a contrast enhancement technique. With a tractable data representation, the necessary enhancement procedures and loss functions are integrated with FSDG-Net. A MID-Net is used to separate the input image into R and T, and a Glow-Net is used to enhance the illumination map in the proposed FSDG-Net. The FSDG-Net that has been proposed is unique not only in that it can learn by correlation consistency of decomposition itself, but also in that it can work regardless of the exact type and even volume of training information. Furthermore, the authors provide the first large-scale dataset for images captured in low light conditions. Extensive experiments conducted in low-light conditions and compare the proposed method with state-of-the-art methods demonstrate its superior performance in terms of subjective and objective evaluations.
Potential study avenues include complex shadow difficulties, stereo vision, and extending the present scenario for significant challenges in high dynamic range. All of them present chances for more investigation.
[2] By applying the updated YOLOv3 algorithm using a fresh data augmentation and labeling approach and taking into account the short dataset's characteristics, Zhang Beini et al. analyzed the DPCR pictures with inconsistent noise and different illumination. The conventional threshold segmentation technique's average precision is less than 65%, and its rate of false positives can exceed 40%, making effective detection difficult for this activity.
When compared to conventional algorithms, deep neural networks have two drawbacks: they require more training samples and take longer to identify data. However, these drawbacks are significantly mitigated by the RBTM and STAM suggested in this paper. Without compromising the dPCR's standout visual characteristics, it may quickly increase the dataset and reduce the labeling time by more than 70%. Additionally, the suggested enhanced YOLOv3 model's average accuracy of 98.98% demonstrates its clear advantage over comparable techniques. With a processing performance that is 1.68 times faster than Mask R-CNN and YOLOv3 separately, and 4.5 times faster overall, the proposed model provides state-of-the-art dPCR image identification.
Additionally, the suggested model is lighter than YOLOv3 and Mask R-CNN, increasing the likelihood that it will be implemented on an embedded system and able to meet the demands of real-time and high-throughput dPCR. The authors want to use this paradigm in the future to create dPCR in real-time on embedded and portable devices.
[3] Yihang Ding et al. recommend the Temporal Relational 530 Cross Transformers according to Picture Difference Pyramid 531 (TRX-IDP) approach for few-shot action identification.
On TRX, the suggested 532 technique is based. Because of this, the frame tuples 533 that were utilized in the query undergo scaling, sigmoid augmentation, and high-order image difference. When paired with Motion 535 History Image (MHI), it creates the Image Difference Pyramid (IDP) 536 that contains information on motion features.
In addition, authors 537 rethink and expand the model's linear mapping function and for IDP 538, create the Cross Transformers query representation. On HMDB51, UCF101, and partial SSv2, the TRX-IDP approach provides state-of-the-art 541 performance; however, on Kinetics-400 and complete 543 SSv2, it falls far short of Hy RSM. On few-shot 540 benchmarks across all four datasets, it outperforms TRX. The authors would like to look into and maybe include other metric-based few-shot action recognition methods 545 in the IDP module 544 in the future.
[4] Min Jun Lee et al. improved the nearest-neighbor classification technique currently in use for few-shot learning by introducing a metric-based training methodology. The authors used mixed, triplet, and cross-entropy losses to train the embedding network, and they analyzed the results for nearest-neighbor classification. Using the triplet loss has a major impact in the

1-shot setting. The same loss showed the best accuracy for the 5-shot setting in the un-normalized configuration, and similar accuracy was displayed for the full setups. However, in order to train on such backbones with constrained GPU memory, the scientists intend to make certain adjustments to their current recommended model.

The recommended triplet loss training model makes heavy use of GPU memory, which hinders the deployment of advanced backbones like DenseNet or ResNet. To improve similarity learning, mixed loss will also be developed. Further research will concentrate on an ensemble model that trains utilizing both triplet loss and cross-entropy loss.

The following parts make up this literature survey study: part II assesses earlier studies on the structure of literature surveys; Part III offers conclusions and suggestions for further research.

2 RELATED WORK

[5] ZhonghuaGuo et al. presented an enhanced approach to the few-shot fish classification problem. With this method, authors may create realistically rendered fake images from a small amount of training data, resulting in a wide diversity of few-shot fish datasets. As a result, it performs better when classifying fish images in a few shots. The findings of the experiment indicate that by using more fictitious photos in training sets, authors can achieve greater classification accuracy.

[6] EmmanouilPatsiouras et al. talked about some real-world issues with how several common data-hungry training approaches train. Some writers have discussed the potential benefits of few-shot learning techniques in removing these obstacles and achieving sufficient generalizations on specific image identification tasks. Authors in the field of unmanned aerial vehicle (UAV) sports cinematography have employed few-shot learning techniques to categorize athletes into distinct groups, each of which represents a subset of the original image class. Authors have suggested a novel method for addressing these kinds of recognition challenges, and they have demonstrated how the novel method can maximize recognition accuracy through suggested experiments.

[7] Sun Yuanshuang and others. This paper presents the creation of an attribute-guided GAN and the proposal of a strengthened training for episodes technique for the few-shot SAR image-generating problem. The authors test the effectiveness of the proposed method using both real and simulated data. The experimental results show how effective the suggested method is, and using different source data for training is made possible by the enhanced episode training strategy. Two different kinds of recognition studies, with just five samples per category, visual inspection, and picture similarity measures all support the improved accuracy of the generated data. The utilization of the proposed AGGAN model's images can result in a minimum 4% increase in the identification rate for the 5-shot SAR target detection challenge.

The experiment's results also demonstrate that, when it comes to the difficulty of generating few-shot SAR photographs, there is still space for improvements in terms of the generated photos' image quality.

For instance, additional attribute data can be utilized to create comparable images by serving as conditional labels. In the proposed study, the authors primarily address the land vehicle target at this time. In order to develop ship targets, further research may look into using the proposed approach to satellite SAR data.It's also necessary to look into how the produced and simulated data are applied.

[8] PengYishu and colleagues. A few-shot learning method based on convolutional transformers is proposed in this study for cross-domain hyperspectral image classification.

Three main parts make up the method to reduce the dimensionality of the data: a few-shot learning-based distribution aligner; a convolutional transformer network-based feature extractor; and a fully convolutional network-based local-global features domain discriminator to minimize domain shift. The usefulness of the suggested CTFSL has been demonstrated through experiments that were carried out on three different genuine hyperspectral images. The results demonstrate that it performs better in cross-domain HSI classification than the current state-of-the-art FSL approaches. Nevertheless, the suggested CTFSL method's high computational cost is necessary for its good performance. This work's performance should be further enhanced while the computation time is decreased with additional advances.

[9] For granular target identification, Yuan Tai et al. presented the CDFSL-DIM method in this research. The CF-contrastive loss is a suggested technique for acquiring unique shared features. Additionally, the authors suggest a unique training method called SQE, which modifies the micro batch's learning rates by the training samples' heteroscedastic uncertainty. For the MSTAR, P-Open SAR ship, and P-VAIS datasets, the maximum accuracy improvement is 30.09%, 7.32%, and 12.92%, respectively, when compared to other meta-learning-based CDFSL methods. For these datasets, the associated accuracy improvement is 11.47%, 4.98%, and 7.00% when compared to transfer-based techniques. Algorithms for lowering the target network's parameter count should be investigated for future development.

[10] The need for massive volumes of data to train ever-larger models is described by M. b. Bijoy et al.; nevertheless, it is getting harder to clean datasets at this size. These massive models' integrity and overall performance are threatened by noisy data points. Dealing with outliers and noise in datasets is essential for writers as they transition to an ecosystem of systems driven by AI. An automated solution to this issue is provided by this study. A cleaner technique that is based on statistics may be more appropriate in situations like medical images, in cases when the dataset's distribution is less spread out. Other computer vision domains that face challenges from noisy input can also benefit from this strategy. It is evident from the aforementioned findings and visualizations that representation learning techniques based on Deep Learning are capable of automating the cleaning process. To improve the model, future work will involve training it using a wider variety of datasets.

[11] A lightweight and easy transductive learning technique for the few-shot image classification challenge, the Few-shot Cosine Transformer (FSCT) was introduced by Quang-Huy Nguyen et al., based on the archetypal network and vision transformer. The authors made two iterations of improvements to the suggested framework's learnable prototypical embedding to strike a balance between easy and challenging samples of the labeled support instances supplied. The first step in few-shot recognition was to use cosine similarity as the foundation for cosine attention to calculate a correlational map between the support and query samples. According to the authors' thorough study and experiments, cosine similarity aids the attention mechanism by providing a more reliable and consistent attention output in the form of a correlational map. As a result, FS-CT can perform competitively across few-shot datasets across a wide range of configurations and settings. Several visual transformer-based few-shot algorithms also show improved performance when suggested cosine attention is applied, according to the empirical results. Lastly, the authors show how FS-CT might be helpful in healthcare research by employing a unique dataset of yoga positions. To further understand how pre-trained models affect few-shot learning algorithms and whether the proposed learnable prototype embedding can handle hard data, additional research is needed.

[12] A novel FRI Net was presented by Qinglong Cao et al. Rotation-adaptive matching is specifically carried out by utilizing support data that differs in orientation but is consistent across categories, to make soaring objects in various orientations active. Meanwhile, complementary rotation-varying segmentation results led by the same ground truth are fused to provide the final rotation-invariant parsing result. Also, the pre-trained backbones in the basis classifications are given freshly to offer a better feature space. Comprehensive testing using the few-shot overhead semantic segmentation benchmark demonstrate that the proposed model achieves state-of-the-art performance.

[13] Jiawei Shi, Zhiguo Jiang, and colleagues introduced the NNPR, an efficient few-shot ship categorization method. The writers train the unified CNN feature extractor separately and then use the closest neighbor technique to get a representation of the prototype. Experiments show that the suggested method outperforms state-of-the-art few-shot deep learning algorithms as well as conventional methods. The outcomes can serve as a standard for few-shot ship categorization and inspire additional research on this difficult issue from other sci-

entists. In subsequent research, the authors plan to enhance the prototype's representation and present a comprehensive method for combining ship detection and few-shot ship classification.

[14] Yunqing Zhao and colleagues presented Born-Again Network (BAN) episodic instruction for domains generalization few-shot classification (DG-FSC). They found that BAN produces stronger decision boundaries and more discriminative features on new tasks from unexplored domains. This suggests that, similar to the observation in standard supervised learning, BAN offers hope for DG-FSC problems. To the best of our knowledge, no one has ever used BAN for episodic training. Because of this, the writers propose Few-Shot BAN (FSBAN) as their main contribution. The multifunctional learning objectives of FS-BAN are Meta-Control of the Temperature, Mismatched Teacher, and Mutual Regularization. Overfitting and domain shift are two separate problems that are highlighted in DG-FSC. They aim to address these concerns. FS-BAN's efficacy is demonstrated by the outcomes of six benchmark datasets, three baseline FSC models, and qualitative and quantitative ablation study results.

[15] Zhang Chengyeet coll. To perform few-shot hyperspectral image classification, this work suggested a GPN for Guided sample size reduction for hyperspectral image classification. The following primary findings were drawn from the tests, which were carried out for verification. In the case of tiny samples, the accuracy of GPN is superior to that of popular approaches now in use. According to the comparison analysis, the GPN is cutting-edge when it comes to utilizing less supervised samples to solve the classification problem of hyperspectral images. According to the ablation study, the global representation learning technique makes the biggest contribution to improving accuracy, however, the dense branch and SSAN branch modules are also helpful. The time needed to put the suggested method (GPN) into practice is comparable to other popular techniques in the same operational setting. The next research should employ more hyperspectral datasets to see whether the proposed GPN is helpful for few-shot hyperspectral picture categorization.

[16] Gege Song et al. discuss how few-shot learning becomes more crucial when real data is lacking. For the few-shot recognition of unidentified restaurant food photographs, the researchers propose a hybrid attention-based prototype network. Convolutional neural networks are initially used to extract information from food images. Next, the class prototype of the prototypical network is obtained using the instance-based attention mechanism, but the model can more efficiently capture feature-level features by using the feature-based attention mechanism. To create a better representation of the image, a feature-based attention mechanism directly integrates visual features into attention weights. The evaluation on the benchmark food picture classification dataset shows that the suggested strategy achieves greater accuracy and beats several state-of-the-art algorithms. Further investigation into hybrid attention-based prototype networks' performance evaluation

, encompassing picture feature extraction and attention mechanism design, will be carried out.

[17] In order to deal with the specific challenges related to few-shot categorization in SAR images, Yipeng Zhang et al. present a new framework named SPT-FSC in this article. Some of the main suggestions include improving the ship classification dataset, designing an RFFA mechanism, building a TEB to get SPT embedding, and inventing SPT. To classify ships in SAR photos, the testing outcomes validate the effectiveness of the suggested SPT-FSC approach. Ship targets' basic structural and shape properties are captured by the SPT, on the other hand, the TEB allows the network to take advantage of the inherent data that every scattering point has.

The RFFA technique recognizes the contributions and reciprocal interactions between features from several sources, which improves the discriminative power of fused features. Furthermore, the network can now more effectively address the few-shot ship classification issue in SAR images because to the creation of an enhanced ship classification dataset.

Numerous assessments shed light on how each element affects the suggested model's overall performance. These tests provide helpful insights into the behavior of the model and validate the significance of each element in improving overall performance.
The proposed SPT-FSC strategy greatly improves few-shot ship classification in SAR images by addressing challenges related to imaging variability, diversity of training data, and the requirement for special modifications to existing few-shot classification methods.
[18] These are Chao Xuewei and his colleagues. According to this study, KNN distance entropy measurement is a good method to analyze image data for semi-supervised few-shot remote sensing image distribution. Applying a meta-task experimental approach and examining the meta-task average accuracy, this study reduces the impact of a random experiment while enhancing the strength of the findings. The results demonstrate that because of the data's rigidity, the primary data are appropriate for small research. We show that KNN's semi-guided data filtering based on distance entropy significantly improves the performance of many experimental domains. To help understand the motivation and context behind the plan, we show the unique distribution of analysis profiles and analyze the impact and profile size of several negative patterns.
[19] Wang Bing and others. In this paper, for the few-shot remote sensing picture scene classification task- TDNet, a unique transductive learning framework, was recommended. In order to produce anisotropic embedding for every pair of the support category and instance, the authors first proposed the conditional metric embedding technique. The model's flexible scalability allowed it to account for the metric biases between classes thanks to its architecture. In order to leverage the help of unlabeled cases for accurate prototype estimate against intraclass changes, the authors also presented a transductive prototype learning technique in their second work. Third, with a long-term consistency regularization, the authors suggested a greater constraint on the prototype's discriminability across prior tasks. Numerous tests confirmed the suggested method's superiority and validity under low-shot circumstances. Concurrently, the suggested techniques can be utilized straightaway for assignments involving new categories without requiring any adjustments, and they considerably enhance classification precision. Finally, the models are intended to continually store the knowledge of previous jobs and efficiently adapt to new ones, taking into account the gradual nature of fresh concepts in the actual world. In order to accomplish effective self-evolution, writers will concentrate on logically mixing few-shot learning and continuous learning in subsequent work.

3 CONCLUSION AND FUTURE SCOPE

The use of AI for illness detection and image analysis is growing in popularity. Retinal glucose analysis, oral cancer, cervical cancer, and other conditions are examples of these conditions. One popular technique for collecting data is using specialized camera modules to take pictures of the impacted areas. This procedure is prone to error, just like any other data source, and it might include regions and objects that are undesirable and need to be removed. These kinds of datasets contain outliers, which can negatively impact machine learning model performance.
Cleaning data by hand would be quite time-consuming, particularly if it came from several sources. Therefore, before training the model, the data must be cleaned. Hence, this paper analyzes many current trending approaches that deal with the cleaning of the few-shot images. And introduce a hybrid neural network model in this work to suggest a few-shot image dataset cleaning process. To begin, the large few-shot dataset is trained using a 2D CNN neural network model. Then, a channel boost CNN is trained on target distributed images to suitably enhance the color channels. The learned models from 2D CNN and CB-CNN are then combined into a hybrid model to determine the optimal way to clean the few-shot images. By combining the two neural networks and hybridizing the models, the proposed strategy will yield positive outcomes which will be seen in the coming edition of this article.

References

[1] R. Khan, Q. Liu and Y. Yang, "A Deep Hybrid Few Shot Divide and Glow Method for Ill-Light Image Enhancement," in IEEE Access, vol. 9, pp. 17767-17778, 2021, doi: 10.1109/ACCESS.2021.3054505.
[2] Z. Beini, C. Xuee, L. Bo and W. Weijia, "A New Few-Shot Learning Method of Digital PCR Image Detection," in IEEE Access, vol. 9, pp. 74446-74453, 2021, doi: 10.1109/ACCESS.2021.3081562.
[3] Y. Ding and Y. Liu, "A Novel Few-Shot Action Recognition Method: Temporal Relational Cross Transformers Based on Image Difference Pyramid," in IEEE Access, vol. 10, pp. 94536-94544, 2022, doi: 10.1109/ACCESS.2022.3204404.
[4] M. J. Lee and J. So, "Metric-Based Learning for Nearest-Neighbor Few-Shot Image Classification," 2021 International Conference on Information Networking (ICOIN), Jeju Island, Korea (South), 2021, pp. 460-464, doi: 10.1109/ICOIN50884.2021.9333850.
[5]Z. Guo et al., "Few-shot Fish Image Generation and Classification," Global Oceans 2020: Singapore – U.S. Gulf Coast, Biloxi, MS, USA, 2020, pp. 1-6, doi: 10.1109/IEEECONF38699.2020.9389005.
[6] E. Patsiouras, A. Tefas and I. Pitas, "Few-shot Image Recognition for UAV Sports Cinematography," 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), Seattle, WA, USA, 2020, pp. 965-969, doi: 10.1109/CVPRW50498.2020.00127.
[7] Y. Sun et al., "Attribute-Guided Generative Adversarial Network With Improved Episode Training Strategy for Few-Shot SAR Image Generation," in IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 16, pp. 1785-1801, 2023, doi: 10.1109/JSTARS.2023.3239633.
[8] Y. Peng, Y. Liu, B. Tu and Y. Zhang, "Convolutional Transformer-Based Few-Shot Learning for Cross-Domain Hyperspectral Image Classification," in IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 16, pp. 1335-1349, 2023, doi: 10.1109/JSTARS.2023.3234302.
[9] Y. Peng, Y. Liu, B. Tu and Y. Zhang, "Convolutional Transformer-Based Few-Shot Learning for Cross-Domain Hyperspectral Image Classification," in IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 16, pp. 1335-1349, 2023, doi: 10.1109/JSTARS.2023.3234302.
[10] M. B. Bijoy et al., "Deep Cleaner—A Few Shot Image Dataset Cleaner Using Supervised Contrastive Learning," in IEEE Access, vol. 11, pp. 18727-18738, 2023, doi: 10.1109/ACCESS.2023.3247500.
[11] Q. -H. Nguyen, C. Q. Nguyen, D. D. Le, and H. H. Pham, "Enhancing Few-Shot Image Classification With Cosine Transformer," in IEEE Access, vol. 11, pp. 79659-79672, 2023, doi: 10.1109/ACCESS.2023.3298299.
[12] Q. Cao, Y. Chen, C. Ma, and X. Yang, "Few-Shot Rotation-Invariant Aerial Image Semantic Segmentation," in IEEE Transactions on Geoscience and Remote Sensing, vol. 62, pp. 1-13, 2024, Art no. 5600513, doi 10.1109/TGRS.2023.3338699.
[13] J. Shi, Z. Jiang, and H. Zhang, "Few-Shot Ship Classification in Optical Remote Sensing Images Using Nearest Neighbor Prototype Representation," in IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 14, pp. 3581-3590, 2021, doi: 10.1109/JSTARS.2021.3066539.
[14] Y. Zhao and N.-M. Cheung, "FS-BAN: Born-Again Networks for Domain Generalization Few-Shot Classification," in IEEE Transactions on Image Processing, vol. 32, pp. 2252-2266, 2023, doi: 10.1109/TIP.2023.3266172.
[15] C. Zhang, J. Yue, and Q. Qin, "Global Prototypical Network for Few-Shot Hyperspectral Image Classification," in IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 13, pp. 4748-4759, 2020, doi: 10.1109/JSTARS.2020.3017544.

[16] G. Song, Z. Tao, X. Huang, G. Cao, W. Liu and L. Yang, "Hybrid Attention-Based Prototypical Network for Unfamiliar Restaurant Food Image Few-Shot Recognition," in IEEE Access, vol. 8, pp. 14893-14900, 2020, doi: 10.1109/ACCESS.2020.2964836.
[17] Y. Zhang, D. Lu, X. Qiu, and F. Li, "Scattering-Point Topology for Few-Shot Ship Classification in SAR Images," in IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 16, pp. 10326-10343, 2023, doi: 10.1109/JSTARS.2023.3328066.
[18] X. Chao and Y. Li, "Semisupervised Few-Shot Remote Sensing Image Classification Based on KNN Distance Entropy," in IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 15, pp. 8798-8805, 2022, doi: 10.1109/JSTARS.2022.3213749.
[19] B. Wang, Z. Wang, X. Sun, Q. He, H. Wang and K. Fu, "TDNet: A Novel Transductive Learning Framework with Conditional Metric Embedding for Few-Shot Remote Sensing Image Scene Classification," in IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 16, pp. 4591-4606, 2023, doi: 10.1109/JSTARS.2023.3263149.