Classification for Synthetic Aperture Radar Ship Image of the Ship Targets under the Few-shot Condition

Siqi Zhang*

Department of Mathematics, The Ohio State University,100 Math Tower, 231 West 18th Avenue Columbus, Ohio, USA 43210-1174

* Corresponding Author Email: Zhang.10074@buckeyemail.osu.edu

Abstract: Deep learning algorithms have shown outstanding results for the categorization of synthetic aperture radar (SAR) images. But because there is a conflict between the deep learning techniques' broad parameter space and the scanty labeled samples of ship targets, most deep learning models are unable to provide satisfying outcomes under the few-shot scenario. Classification for SAR ship images often faces the few-shot condition because they are sensitive to system parameters, so that the amount of effective SAR images is very low. To this end, this article proposes a method based on graph neural network (GNN). An end-to-end classification model was trained by this method. Moreover, our method extracts multi-scale features of the targets to make classification effect better. The dropout layers are added into the graph convolution layer to reduce the overfitting in few-shot classification. Finally, our model is applied to the mixed dataset of the simulated and real data to verify effectiveness of our method. The experimental findings presented in this research show that our method outperforms other cutting-edge approaches in terms of categorization rate.

Keywords: Graph Neural Network (GNN), Synthetic Aperture Radar (SAR), Few-shot learning, Ship target.

1. Introduction

A microwave detection tool that is active is called Synthetic Aperture Radar (SAR), which could complete the land detection and high-resolution imaging well under bad weather like snow, rain, sandstorm etc. Synthetic aperture radar (SAR) images, which have been used as a necessary data source for monitoring maritime activities, have a wide range of uses, for example, the oil and ship detection application has been the focus of many previous studies. The classical SAR target classification recognition like constant false alarm rate method [1], getting the target parameter via hand-designed features, for example, geometric features, scattering center to detect ship targets. For the view of imaging, the radar beam light the area at a certain depression angle, so that the side-view SAR image has inherent characteristics such as shadow, top-bottom inversion etc., resulting in a small number of effective samples for SAR images. Those problems have a negative impact on SAR ship detection target recognition. Moreover, speckle noise in SAR images would makes the ship target in the image exists in the form of a distributed scattering center, which makes the imaging sidelobe of the strong point target obvious. That is because ship targets occupy fewer resolution units under low resolution due to speckle noise. Furthermore, targets in SAR images are sensitive to system parameters, especially attitude angles. Therefore, target recognition become hard of ships.

To be able to represent complex functions of high-level abstract concepts numbers, to solve AI-related tasks such as object recognition, voice, and language, requires deep learning [2]. Deep learning methods try to find the internal structure of the data and discover the relationship between variables. Numerous studies have shown that the way data is represented has a significant impact on the success of training learning. It can eliminate the influence of changes in the input data that are not related to the learning task on the learning performance, while preserving the use information of learning task. Convolutional neural network (CNN) and graph neural network (GNN) models are two types of small sample image categorization techniques that may be grouped together. The algorithm includes four learning paradigms: transfer learning, meta-learning, dual learning, and Bayesian

learning. Meta-learning analysis was first proposed by Gene V. Glass in 1976[3], which analysis many statistical analyses, using a secondary analysis method.

To apply CNN to radiological problems, short datasets, and overfitting, Yamashita et al. [4] provided two examples. This paper will cover CNN, small dataset and solving method to minimize. This work suggests a small sample image recognition approach based on CNN is by using convolutional neural network as the feature extractor of images, which could improve the accuracy of image recognition with only a few training samples. Few-shot Learning merely contains a limited number of supervised data information, which is a type of machine learning problems. By using few-shot learning, it could learn enough information about a category from one or just several of images. Zhen Tan et al (2022 tried to optimize the baseline model of Incremental Learning paradigm by creating a Hierarchical learning method [5], which worked by calculating the task sensitive regulation value to decide whether need to add additional nodes on the Hierarchical method. In the same year, Zhirui Zhang et al has proved that the few-shot learning could solve two major problems in classifying Galaxy Morphologies, which are extensive training sets and potentially biased outcomes [6]. As a result, they employed supervised deep learning and few-shot learning linked to Siamese networks to classify the data set, and they get that the few-shot learning has a better accuracy rate in classifying Galaxy Morphologies. The study of few-shot learning is basically separate from three parts, data, model, and algorithms. The methods of few-shot learning can be roughly separated into data augmentation-based, model-based optimization, and Algorithmic Optimization-Based Method. Model-based optimization includes embedding learning (use in this article), which mapped samples to a low-dimensional space, thus achieving the effect of reducing the hypothesis space.

By using message transmission between graph nodes, graph neural networks can represent the reliance of graphs. Graph neural network approaches are used in recommender systems by Shiwen Wu et al because to its superiority in graph representation learning and the fact that much of the information in recommender systems fundamentally has a graph structure [7]. Thanks to the success of CNN in multiple domains, many re-define creating approaches that parallel convolution theory for graph data. These methods fall under the category of convolutional graph neural networks (ConvGNNs) [8]. Spectral-based techniques and spatial-based methods make up the two main groups of ConvGNNs. This paper will use spatial-based ConvGNNs, propagating node information along edges is the nature of the spatial graph convolutional operation. The entire SAR picture is treated as a graph node.

In this study, the author implements few-shot categorization of SAR ship photos using an end-to-end graph neural network. To aid in the categorization of SAR pictures, a multi-scale feature extraction technique is proposed. Finally, drop-out layers are introduced into the network to reduce the overfitting.

2. Dataset

This article totally uses 15 categories of ship targets, 7 of them are simulated data, 8 of them are from OpenSAR [9], which is an open source for SAR shipment images. Only 30 images are simulated in each class to satisfy the condition of small samples. Each image with size 128*128 is resized to patches with size 32*32, 64*64 and 128*128. In test process, eight images are randomly selected for each class. Figure 1 gives part of the images in this article.

The machine used for all the testing has an AMD Ryzen 5 2600X six-core processor clocked at 3.60 GHz and an NVIDIA GeForce GTX 1650 SUPER GPU. The author trained with a learning rate 0.001 and 100,000 iterations, using Xavier method to initialize the weights. Every time the author calls 5 shots, every shot has 20 images, the batch size is 12. Each method will repeat ten times.

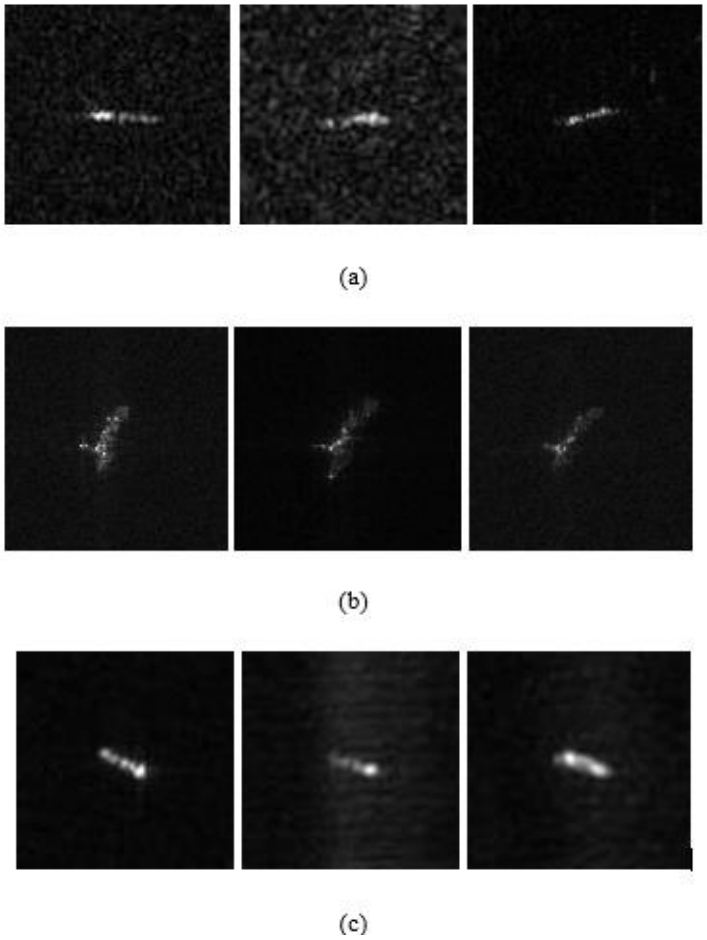


Figure 1. Images of 3 classes in the dataset

3. Algorithms

3.1. Difficulties of SAR ship target classification

As the key targets of maritime monitoring and wartime strikes, ship targets are a crucial component of remote sensing picture automated target detection and recognition. Researchers have conducted a lot of research on the automatic detection and recognition of ship targets in SAR images, mainly using ship targets. The variation in the ship target's electromagnetic scattering properties compared to the sea body in the SAR picture [10].

Classical SAR image ship classification have 3 steps, preprocessing, feature extraction and classification recognition. The difficulties mainly lie in feature extraction and feature selection. On the one hand, manual design and selection of features is a time-consuming and labor-intensive method, and whether it can be selected well depends on experience and luck. On the other hand, the artificially designed features have limitations and cannot fully reflect the high-level information contained in the image. The ship targets on the sea have three-dimensional rotation because of the effect of bad environment at sea. Meanwhile, due to the working modes of SAR, which is multi-angle, multi-frequency, and various working modes, the SAR imaging characteristics of same ship target has different attitudes, stern directions, and motion directions. The probability of the target being observed, especially some high-value SAR imaging characteristics is small. Many small samples in the SAR imaging of the ship target makes it difficult to use the existing deep learning network to identify.

In addition, the following characteristics SAR image ship target have will also increase the difficulty of classification:

(1) From the perspective of SAR imaging method, the radar beam illuminates the area to be mapped at a certain depression angle, so that the side-view SAR image has inherent characteristics such as shadow, perspective shrinkage, and top-bottom inversion, resulting in a small number of effective samples of the SAR image obtained. When the SAR separation rate increases, the above-mentioned problems will occur in ship targets which will have a certain impact on the fine-grained classification of the SAR ship target.
(2) The imaging difference is highly different and the target in the SAR picture is sensitive to the system parameters, particularly the attitude angle. Therefore, the difficulty of target invariant feature extraction increases the difficulty of target recognition.
(3) The feature extraction of ship radiated noise has always been a difficult point for studying the problem of ship target recognition. The reason is the complexity of the marine environment and the particularity of the underwater acoustic channel. Those make it is difficult to obtain ship signal features from the complex background noise

3.2. Graph Neural Network (GNN)

To do few-shot learning's target categorization, the author might utilize a graph neural network. Joan and Bruna et al propose using graph neural network (GNN) to do transfer learning [11]. Consider input-output pairs $\left(\mathcal{T}_i, \mathcal{Y}_i\right)_i$, the author could separate them into three parts, as shown in in equation (1).

$$T = \left\{\{(x_1, l_1), \ldots (x_s, l_s)\}, \{\widetilde{x_1}, \ldots \widetilde{x_r}\}, \{\overline{x_1}, \overline{x_t}\};\ l_i \in \{1, K\}, x_i, \widetilde{x_j}, \overline{x_j} \sim P_i(R^N)\right\}\ Y = (y_1, \ldots y_t) \in \{1, K\}^t \quad (1)$$

Where s denotes the quantity of samples that have labels, r denotes the quantity of unlabeled samples (r > 0 for the semi-supervised and active learning situations), and t is the quantity of samples that need to be classified. K is the number of classes. The author could let t=1 when only using one sample to verify in each mission. P is a collection of images that follow a given category classification. When r=0, t=1, s=q*K, only one image in the collection has an unknown label. In this paper, the author only uses labeled samples which appears q times, it can be considered as a q-shot, n-way few-shot learning.

By focus on the case t = 1, the author just classifies one sample per task T. $P_l(\mathbb{R}^N)$ express a class-specific image distribution over $\mathbb{R}^N$. In our paper, the targets Y_i are associated with image categories of designated images $\bar{x}_1, \ldots, \bar{x}_t \in T_i$ with no observed label. Under knowing training set $\left\{\left(T_i, Y_i\right)_i\right\}_{i \le L}$, the author considers the standard supervised learning objective using the model Φ (T; Θ) = p (Y | T) and R is a standard regularization objective.

$$\min_{\theta}\left(\frac{1}{L}\right)\sum_{i \le L} l(\emptyset(T_i; \theta), Y_i) + R(\theta) \quad (2)$$

GNN can be used to achieve the target classification for few-shot learning. The GNN differs from the conventional convolutional neural network in that it transforms the input data into graph data in a non-Euclidean space. The pixels in picture data, for example, are placed sequentially and consistently in Euclidean space. However, data of non-Euclidean space, such as Graph data, the arrangement rules are disordered and not clear. The good point is the author do not care how its data points are arranged because it has no use for data analysis, the connection between nodes and nodes is more useful for us. One of the structures in GNN is the graph convolutional network (GCN), which may input an entire picture as a node or utilize each pixel in the image data as a node in the graph. The following are crucial networking concepts: $G = (V, E, A)$ where A denotes the proximity matrix, V denotes the set of nodes, and E denotes the set of edges. $v_i \in V$ describes a node, $e_{ij} = \left(v_i, v_j\right) \in E$ where V stands for the set of nodes, A for the proximity matrix, and E for the set of edges. A is a N*N matrix in this case.

$$A = \begin{bmatrix} 0 & w_{12} & w_{13} & \cdots \\ w_{21} & 0 & w_{23} & \cdots \\ w_{31} & w_{32} & \ddots & \cdots \\ \vdots & \vdots & \vdots & 0 \end{bmatrix} \tag{3}$$

In equation (3), A is the matrix which all the elements in diagonal are 0, and it is a symmetrical matrix $w_{ab} = w_{ba}$. Here, $A = MLP\left(abs\left| f_a - f_b \right| \right)$, where f_a means the corresponded a_{th} node feature matrix, MLP (Multilayer Perceptron). Figure 2 intuitively shows the steps of learning node representations for graph neural networks:

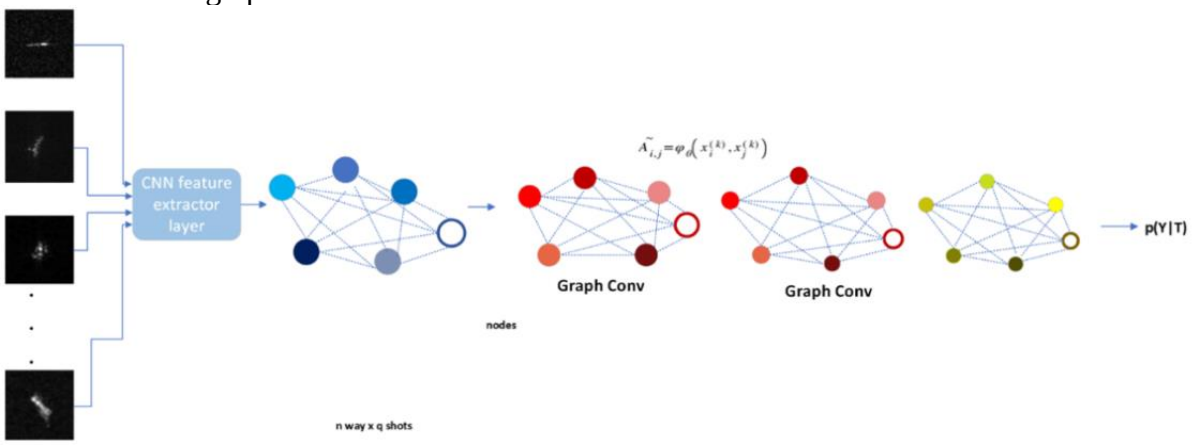


Figure 2. Graph Neural Network (GNN) illustration

At first, images combine with one-hot code through CNN feature extractor layer, they become the feature of nodes (blue). The different colors of circles are represented for different nodes features, and lines are represented for the relation of nodes which could be represented by adjacency matrix numerously. The big circle represents a series of nodes. The relationship between the nodes is updated through the graph convolutional network layer. Finally, the distribution p(Y|T) is calculated, output the predictive analytics outcome under known labels. The author could use $W\left(x_{new} \right) = A \cdot x_{old}$ to explain graph convolution network. A is the proximity matrix, x_{old} is the old nodes, $W\left(x_{new} \right)$ is linear layer. A multiply the old nodes become new linear layer. The form under describes the whole process of the algorithm in this article including the input, output, and algorithm. Overall network structure is shown in Figure 3.

Table 1. Algorithm of the multiscale GNN for SAR images classification

Algorithm: Multiscale GNN for SAR images classification
Input: SAR image dataset, batch size E
Output: Predict label for input image
Algorithm:
1. Preprocess data
2. Separate the dataset into train and test datasets.
3. Create an end-to-end GNN framework
4. For e=1 to E do
5. Resize images to a*a, b*b, c*c
6. Use feature extraction layer to extract feature F1, F2, F3 of
7. the above images
8. Combined feature F= [F1, F2, F3]
9. Put features into network
10. Calculate loss function L
11. end
12. Conduct label prediction

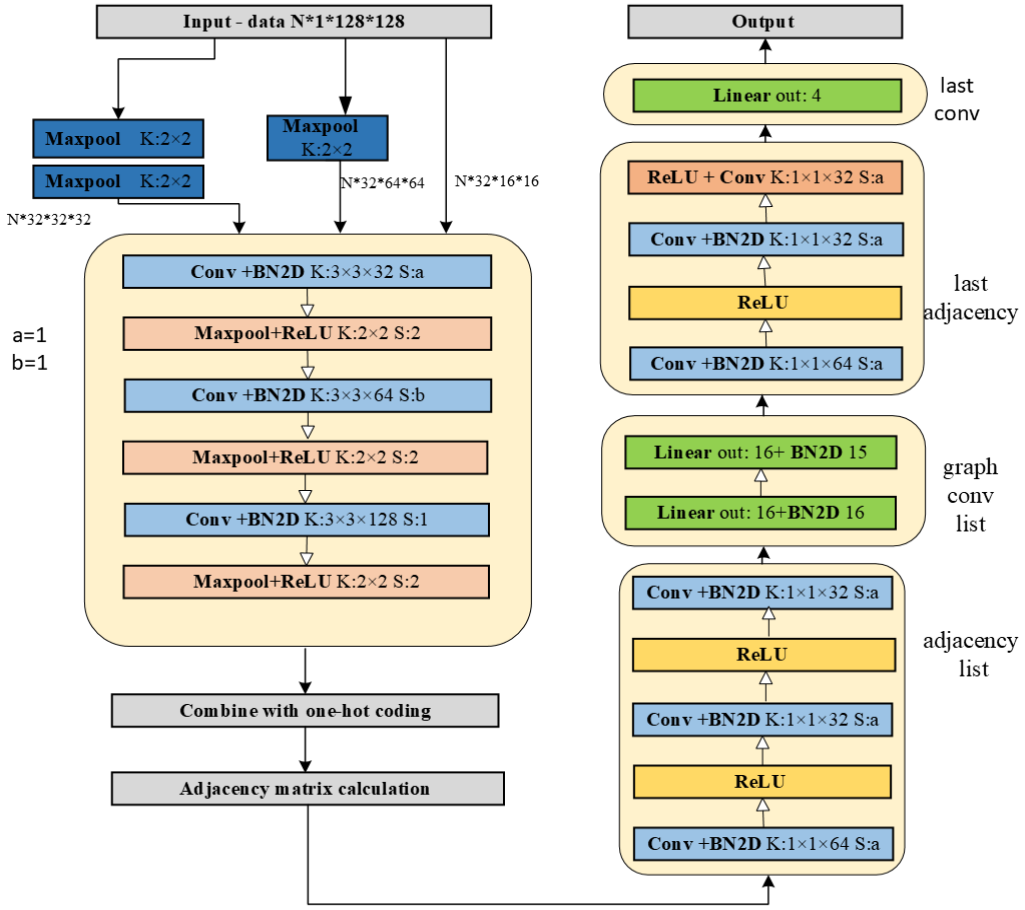


Figure 3. Network structure

4. Results

The network mode is 4 way-20 shot using single scale and multi scale in our paper. Both single scale model and multi scale model have a good accuracy, since the author trained data under few-shot condition, with the value 75.83±0.83 and 79.17±0.83 separately. In the confusion matrices in table 2 and table 3 for multiscale and sing scale, it indicates that multi scale have more correctly predicted labels. The loss of both models is shown in Figure 4 and the author could see our proposed method have better converge speed and lower final loss than the single one.

Table 2. Single scale confusion matrix

		Label			
		1	2	3	4
Prediction	1	73.33%	33.33%	3.33%	0
	2	18.67%	20.00%	20.00%	2.66%
	3	8.00%	33.33%	76.67%	0
	4	0%	13.33%	0	97.34%

Table 3. Multi scale confusion matrix

		Label			
		1	2	3	4
Prediction	1	76.67%	43.33%	6.67%	0
	2	20.00%	26.67%	6.67%	0
	3	3.33%	23.33%	86.67%	0
	4	0	6.67%	0	100%

Table 2 and 3 show the confusion matrix which have reference and prediction accuracy for each category. The number 1 to 4 in both rows and columns represent 4 different classes the paper used. Since our model is 4-way 20 shot model, the author classified data into 4 categories here. The diagonal is the probability of correct prediction, which means that a category classification should be in this class and the prediction is correct. For example, the value in row 1, column 2 represents the classification rate that the class 2 is classified to class 1 by mistake. The sum of each column would be 100% in total. From the comparison of these two tables, it is obvious that class 2 has the lowest classification rate, that is because class 2 is similar as the other categories, so the effect of classification is not good as others. In general, multi-scale has a higher accuracy since the values in diagonal are bigger. But most of classes have a good performance at classification which prove the effectiveness of multi-scale graph neural networks. The average accuracy of class 1, 3 and 4 are around 90%.

5. Discussion

Figure 4 shows the loss curve of both multi-scale and single-scale training. The red graph is single scale feature extraction method applied in GNN and the blue one is innovative method as multi scale which applied to propose network. It could be seen that their loss is both converge to 0.4 after 100,000 iterations. However, the multi-scale method converges quicker than single scale at first and tend to be more stable. Moreover, the blue line has a lower loss than the red. The figure explains that the multi-scale method is helpful to few-shot recognition, and it is more suitable for network training which could describe feature correctly.

Table 4. Classification rate comparison

Method\shot	20 shot (%)
Fine-tuned Alexnet	70.00±0.83
MAML	72.50±0.83
Siamese Nets	74.17±1.66
Single scale	75.83±0.83
Matching Nets	76.67±0.83
Proposed method	79.17±0.83

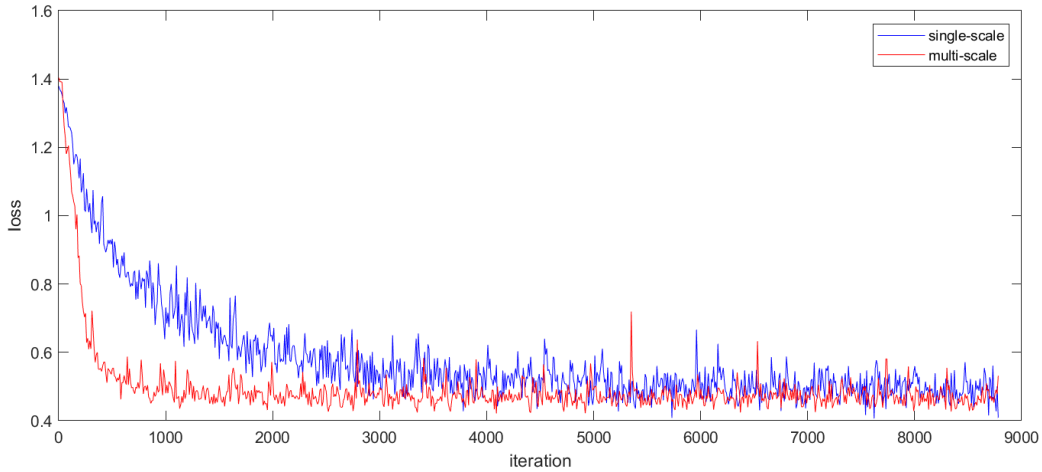


Figure 4. Loss function for single-scale and multi-scale methods

To see the advantage of graph convolution network, from Table 4, this work uses Fine-tuned Alexnet, MAML, Siamese Nets, Matching Nets to compare with our proposed method. Where the ± (minus or add) means the fluctuation of lots of experiments. The outcome is showed clearly that the proposed methods which is multi scale have the highest classification rate.

6. Conclusion

SAR have better function than optical sensor in detecting the ship targets because it has advantages like in bad weather like cloudy, SAR is more visible than traditional optical sensor. However, it is sensitive to the imager parameter which lead to few-shot conditions. This paper trained an end-to-end graph neural network to realize few-shot classification of SAR ship images and use a multi-scale feature extraction method to contribute. Moreover, drop out layers are added into the proposed network to reduce the overfitting. By using both adjacency layer and graph convolution layer, the relationship between nodes is explored to realize few-shot classification of images. For the experimental part, a mixed dataset of simulated and real data is used to train the network. It is demonstrated via comparison experiments that the suggested technique has better accuracy. In the future, the author plan to reduce the parameter of the network so that could let the network train and test faster under keeping the accuracy.

References

[1] Snell, J., Swersky, K., and Zemel, R. "Prototypical networks for few-shot learning." Advances in neural information processing systems 30 (2017).

[2] Erhan, D., Courville, A., and Bengio, Y. "Understanding representations learned in deep architectures." (2010).

[3] Glass, G. V. "Primary, secondary, and meta-analysis of research." Educational researcher 5.10 (1976): 3-8.

[4] Yamashita, R., Nishio, M., Do, R. K. G., and Togashi, K. "Convolutional neural networks: an overview and application in radiology." Insights into imaging 9.4 (2018): 611-629

[5] Tan, Z., Ding, K., Guo, R., and Liu, H. G "Graph few-shot class-incremental learning." Proceedings of the Fifteenth ACM International Conference on Web Search and Data Mining. 2022.

[6] Zhang, Z., Zou, Z., Li, N., and Chen, Y. "Classifying Galaxy Morphologies with Few-Shot Learning." Research in Astronomy and Astrophysics 22.5 (2022): 055002.

[7] Wu, S., Sun, F., Zhang, W., Xie, X., and Cui, B. "Graph neural networks in recommender systems: a survey." ACM Computing Surveys (CSUR) (2020)

[8] Wu, Z., Pan, S., Chen, F., Long, G., Zhang, C., and Yu, P. S. "A comprehensive survey on graph neural networks." IEEE transactions on neural networks and learning systems 32.1 (2020): 4-24.

[9] Huang, L., Liu, B., Li, B., Guo, W., Yu, W., Zhang, Z., and Yu, W. "OpenSARShip: A dataset dedicated to Sentinel-1 ship interpretation." IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing 11.1 (2017): 195-208.

[10] Wang, Y., Ma, L., and Tian, Y. "State-of-the-art of ship detection and recognition in optical remotely sensed imagery." Acta Automatica Sinica 37.9 (2011): 1029-1039.

[11] Garcia, V., and Bruna, J. "Few-shot learning with graph neural networks." arXiv preprint arXiv: 1711.04043 (2017).