Few-shot Learning on Histopathology Image Classification

Ankit Kumar Titoriya
Department of Computer Science & Engineering
National Institute of Technology Patna
Patna, India
ankitt.ph21.cs@nitp.ac.in

Maheshwari Prasad Singh
Department of Computer Science & Engineering
National Institute of Technology Patna
Patna, India
mps@nitp.ac.in

***Abstract*— Cancer detection in histopathology slides is not easy even today. CNN (Convolutional Neural Network) based Object identification and segmentation algorithms work very well. A large dataset of medical images is required for classification which may not be available especially for rare diseases. Therefore, deep learning and machine learning may not be effective for rare disease classification. If CNN architecture is trained on one dataset then it performs well but the same architecture may not achieve good accuracy on other datasets. So, generalization is one of main issues. This paper proposes FSL (Few-Shot Learning) to solve generalization and size of dataset. This paper uses Prototypical networks and MAML (Model Agnostic Meta Learning) simultaneously on four different datasets. Along with this, it has also been checked whether these two networks meet the concept of generalization or not. The paper also finds accuracy of both networks in 2-way, 3-way, and 5-way modes. Simulation results show that MAML achieves accuracy of 84.56% in the 2-way 2-shot 2 query mode. Further, simulation results show that Prototypical Network achieves accuracy of 74.575%, 61.9889% and 45.762% in 2-way 2-shot 2 query mode, 3-way 3-shot 3-query mode and 5-way 5-shot 5-query mode, respectively.**

Keywords— Few-shot Learning, Prototypical Networks, Model-Agnostic Meta-Learning, Histopathology, Cancer, Medical Imaging

I. INTRODUCTION

Cancer causes a lot of deaths in the world. It is a type of disease in which the cells of the body start growing in an uncontrollable manner. If it is found early, the right treatment can be given and the patient has a better chance of staying alive. Otherwise, the situation can become quite dire. According to a WHO study [1], cancer caused nearly 10 million deaths worldwide in 2020. Its cases are increasing day by day all over the world, which means the probability of death of a person is increasing. Histopathology is a method to detect and classify cancer. By using this, a doctor can determine the grade of cancer and plan the strategy for the operation. But still, there is a lack of experienced pathologists in the area of histopathology. With this method, the diagnosis of cancer is extremely difficult. Without experience, no one can make a reliable guess.

Since the invention of DL (Deep Learning) and ML (Machine Learning), there are many techniques and CNN architectures that can recognise and segment images. These techniques provide efficient results. There are also many experiments in medical imaging that use DL and ML on histopathology images. All these methods are data-driven, requiring large amounts of labelled data. It is a very time-consuming task for the pathologist to scan and label medical images from the slides. To avoid this problem, some researchers employ data augmentation so that the model can get a variety of data points to learn from. In the second method, a researcher first trains the model on a large dataset and then fine-tunes it on a smaller dataset. We can say that both strategies use a large amount of data or expend it through augmentation. Another challenging issue of DL techniques is high computation, in which models need to be run on high computational resources.

This study employs FSL techniques that don't necessitate a large-scale dataset. Human intelligence serves as the inspiration for the FSL concept. Even if a small child begins to recognise things after a few exposures, then artificial intelligence methods should not need such a large dataset. Similar to this, FSL learns from a small number of images. This paper employs the prototypical network and MAML on the four different datasets, namely (i) BRACS (BReAst Carcinoma Subtyping) (ii) BrakeHIS (Breast Cancer Histopathological Database) (iii) BACH (BreAst Cancer Histology Images) and (iv) LC25000 (Lung and Colon Cancer). Most of this paper is about the idea of generalization, which is a feature of FSL. FSL methods require a dataset with a high number of classes and a lower number of samples. Thus, this study utilises one dataset for training and another for validation, as well as one for train-validation and another for testing.

This study begins with an overview of the related works in Section 2. Section 3 gives an overview of the methodology. This paper provides a brief overview of the datasets in Section 4. This study highlights the outcomes of experiments described in brief in Section 5. The paper reaches its conclusion in Section 6.

979-8-3503-2028-2/22/$31.00 ©2022 IEEE
DOI 10.1109/CSCI58124.2022.00149

II. Related Works

Cancer is a common disease these days. The majority of cancer diseases occur in old people. According to a report [2], slightly more than 70% of cancer cases include those over 50. In medical imaging, researchers have been trying to develop artificial intelligence for a long time. In 1955, Lee Lusted [3] was the first person to say that computers could be used to diagnose diseases. Lodwick et al. [4] digitized chest X-rays for the first time to make CAD systems, which they used to find lung cancer eight years later. One of the earliest and most extensively researched CAD applications in the 1970s and 1980s included the diagnosis of lung cancer by employing chest radiography images. However, the advent of Deep Learning methods has fundamentally altered this field. Researchers are using different ways to learn, like deep learning, to diagnose cancer.

In [6], M. Nishio et al. suggest a computer-aided diagnosis system. To classify lung cancer images, the study uses homology-based feature extraction with ML algorithms. In [7], E. W. Teh et al. transfer the features extracted from one weakly labelled dataset to a less labelled dataset. The study employs the ProxyNCA method for classification. In [8], Y. M. Attias et al. show the effectiveness of metric learning with proxies in DL. It reduces the training time and uses fewer resources. In [9], A. Medela et al. demonstrate the histopathology-based Siamese network for FSL. This study also shows the transfer of knowledge from one dataset to another. In [10], A. Foucart et al. show artifact identification in histopathology slides using DL methods. The study uses the weak and noisy data for the supervision of CNN. The paper also shows the generalization capability of the network. In [11], N. N. Shaikh et al. employ FSL for artifact identification in histopathology slides. The study uses an approach based on a prototypical network.

U.K. Das et al. use metric-based learning for segmentation in [17]. The study uses CNN with K-Means clustering and median filtering. M. Masud et al. propose a new method for DL-based classification in [18]. The paper uses two transformations for feature extraction (2D Fourier transformation and 2D Wavelet transformation) and then combines the features for classification via CNN. In [19], M. Toğaçar employs the DarkNet-19 with Support Vector Machine (SVM) and the Manta Ray Forging optimization algorithm for classification. In [20], M. Ali et al. suggest a multi-input dual stream capsule network. The network uses convolutional layers' strong feature learning abilities to categorize histopathology images. In [21], S. Garg et al. use CNN with feature extraction methods. The paper shows class activation and saliency maps by using SmoothGrad and GradCAM. In [22], M. Phankokkruad uses ResNet50V2, VGG16, and DenseNet201 with ensemble transfer learning to classify lung cancer. In [23], J. Lin et al. propose Pyramidal Deep-Broad Learning (PDBL), which is a plug-and-play module. PDBL is a backbone module that helps to train CNN with less training time. In [24], R. D. Mohalder et al. use deep learning-based methods to predict colon cancer. In [25], J. Fan et al. propose a transfer learning architecture based on SVM. K. Adu et al. [26] propose a dual horizontal squash capsule network (DHS-CapsNet) for the classification of lung and colon cancers. DHS-CapsNet is a combination of a horizontal squash function and an encoder feature fusion module.

III. Methods

A. Few-shot Learning

In the FSL, there are two types of sets: a support set (S) and a query set (Q). An M-way N-shot classification is a structure of support set in which M stands for M-classes and N for N-samples per class. The Query set contains images drawn at random from N classes. The images in the query set are different from the support set images. Suppose a dataset A contains some E samples with F classes like:

$$A = \{(x_1, y_1), (x_2, y_2), \dots \dots (x_E, y_E) \dots \dots (x_E, y_F)\} \quad (1)$$

Here, $y_i \in \{1,2, \dots . F\}$ is the corresponding label for the sample x. These support and query sets are the subset of superset A where $M < K$ and $N < F$. Here, $x_i \in \mathbb{R}^D$ is a D-dimensional feature vector.

B. Prototypical Network [12]

The prototypical network calculates the prototype $c_f \in \mathbb{R}^g$ for each class using an embedding function $f_\theta: \mathbb{R}^D \to \mathbb{R}^g$ with θ learnable parameters. A prototype for its class is:

$$c_f = \frac{1}{|A_f|} \sum_{(x_i, y_i) \in A_f} f_\theta(x_i) \quad (2)$$

Here A_f stands for the collection of samples tagged with the class f.

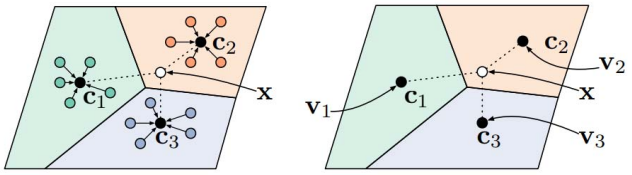

Fig 1: Prototypical networks in few-shot and zero-shot scenarios [12]

Using the distance function $d: \mathbb{R}^g \times \mathbb{R}^g \to [0, +\infty)$, prototypical Networks compute a probability distribution over classes for a sample x using a SoftMax function covering distances towards the prototypes throughout the embeddings:

$$p_\theta(y = f|x) = \frac{\exp\left(-d(f_\theta(x), c_f)\right)}{\sum_{f'} \exp\left(-d(f_\theta(x), c_{f'})\right)} \quad (3)$$

C. Model Agnostic Meta-Learning [13]

The MAML algorithm tries to train a model with quick adaptation to a problem in FSL. Consider a task $T = \{L(x_1, y_1, \dots \dots, x_H, y_H), q(x_1), q(x_{t+1}|x_t, y_t), H\}$ that has the following components: a loss function L, an episode length H, a distribution over beginning observations $q(x_1)$, and a distribution over transitions $q(x_{t+1}|x_t, y_t)$. The algorithm trains the modal to a distribution over tasks $p(T)$. For N-shot learning a task T_i from $p(T)$ learns using N samples through observation

q_i and loss function L_{T_i}. In training phase, model learns from N samples with loss L_{T_i}, and then tested on the new samples of T_i. Then, the model f is refined by looking at how the test error on new samples from q_i changes as the parameters change.

The algorithm's objective is to develop a model that employs gradient-based learning rule to learn fast and no overfit on incoming tasks T_i from distribution $p(T)$. The approach discovers parameters that are responsive to changes in the task. It means that even small changes to the parameters can have a big effect on the loss function for any task. Suppose a model function $f_\theta$with parameters θ. For a new task T_i, parameter $\theta_i'$is computed using gradient descent as shown in equation 4.

$$\theta_i' = \theta - \alpha \nabla_\theta L_{T_i}(f_\theta) \quad (4)$$

Here α stands for step size, which is fixed. In training, parameters of model $f_{\theta_i'}$ optimize for the tasks through distribution $p(T)$ using equation 5.

$$min_\theta \sum_{T_i \sim p(T)} L_{T_i}(f_{\theta_i'}) = \sum_{T_i \sim p(T)} L_{T_i}(f_{\theta - \alpha \nabla_\theta L_{T_i}(f_\theta)}) \quad (5)$$

The algorithms optimise the tasks overall using stochastic gradient descent (SGD). Then equation 6 updates the model parameters as follows:

$$\theta \leftarrow \theta - \beta \nabla_\theta \sum_{T_i \sim p(T)} L_{T_i}\left(f_{\theta_i'}\right) \quad (6)$$

Here β stands for meta step size.

IV. Datasets

This study utilizes four publicly available datasets for the implementation of FSL. In FSL, a model learns from various classes, but medical imaging datasets normally contain fewer classes and samples. Therefore, the study effectively integrates transfer learning with FSL.

A. *BRACS (BReAst Carcinoma Subtyping)*[15]

This dataset was generated through the collaboration of the National Cancer Institute (Italy), the Institute for High Performance Computing and Networking of the Research Council of Italy (Italy), IBM Research (Zurich), and the University of Naples(Italy). The BRACS dataset includes 547 Whole Slide Images (WSIs) that were taken from 189 distinct patients. It also contains 4539 patched images that were taken from 387 WSIs that were gathered from 151 patients. The dataset is made up of seven different groups, namely : (i)Normal Tissue(N), (ii)Pathological Benign (PB), (iii)Usual Ductal Hyperplasia (UDH), (iv)Flat Epithelial Atypia (FEA), (v) Atypical Ductal Hyperplasia (ADH), (vi) Ductal Carcinoma in Situ (DCIS), and (vii) Invasive Carcinoma (IC). The researchers used a 40X magnification factor to capture images from WSI. Table 1 displays the distribution of the dataset.

TABLE I. BRACS Dataset

Data	N	PB	UDH	FEA	ADH	DCIS	IC
Images	484	836	517	756	507	790	649

B. *BrakeHIS (Breast Cancer Histopathological Database)*[14]

This dataset was developed by the P&D Laboratory, Brazil. It contains 7909 histopathology images that were taken from 82 patients. It has eight distinct classes, namely: (i) Adenosis (A), (ii) Fibroadenoma (F), (iii) Phyllodes Tumor (PT), (iv) Tubular Adenoma (TA), (v) Ductal Carcinoma (DC), (vi) Lobular Carcinoma (LC), (vii) Mucinous Carcinoma (MC), and (viii) Papillary Carcinoma (PC). The authors used four magnification factors (40x, 100x, 200x, and 400x) to capture images from WSIs. Table 2 shows the distribution of dataset.

TABLE II. BrakeHIS Dataset

	A	F	PT	TA	DC	LC	MC	PC
40x	114	253	109	149	864	156	205	145
100x	113	260	121	150	903	170	222	142
200x	111	264	108	140	896	163	196	135
400x	106	237	115	130	788	137	169	138

C. *BACH (BreAst Cancer Histology Images)*[16]

This dataset consists of microscopic histopathology images that have been expertly annotated by two pathologists from the Institute for Research and Innovation in Health and the Institute of Molecular Pathology and Immunology of the University of Porto. It was created for the ICIAR 2018 grand challenge. It contains 400 training and 100 testing microscopy histopathology images acquired from 33 patients. This dataset has four groups: (i) Normal (N), (ii) Benign (B), (iii) In Situ Carcinoma (ISC), and (iv) Invasive Carcinoma (IC). Table 3 displays the distribution of the dataset.

TABLE III. BACH Dataset

Data	N	B	ISC	IC
Training Images	100	100	100	100
Test Images	25	25	25	25

D. *LC25000 (Lung and Colon Cancer)*[5]

This dataset was created with the help of James A. Haley Veterans' Hospital. It contains 25,000 histology images with five classes. The dataset is a combination of lung cancer tissues and colon cancer tissues. It consists of five classes, named as: (i) Benign Lung (BL), (ii) Lung Adenocarcinomas (LA), (iii) Lung Squamous Cell Carcinomas (LSC), (iv) Benign Colon (BC) and (v) Colon Adenocarcinomas (CA). Table 4 shows the distribution of the dataset.

TABLE IV. LC25000 Dataset

Data	BL	LA	LSC	BC	CA
Images	5000	5000	5000	5000	5000

V. Experiment And Results on Prototypical Networks and MAML

This study employs four different datasets combinedly on prototypical networks and MAML. This is because FSL focuses on generalisation and learning from fewer samples. Fig. 2. shows the workings of FSL in which there are two types of sets of images: one is a support image set and the other is a query image set. A few-shot learner trains itself using support images. Assume a model is learning using the 5-way, 5-shot method. Here, this means that the support image set will contain a total of 25 images (5 images per class). The query image set may contain any number of images (e.g., 5, 10, or 15). The FSL model learns through the support set and its weights are updated through the loss function calculated on the query image. This process is repeated many times. In every iteration, the model picks the images randomly from the dataset. Assume there are 1000 classes in a dataset and each class contains 50 images. It means that each class has fewer data samples for training a deep learning architecture. The FSL model can easily learn from this and achieve good accuracy. For this, there is a need to create class-wise partitions of the dataset (e.g., 500 classes for training, 300 classes for validation, and 200 classes for testing). Then, the model needs to train itself on the M-way N-shot method, and after that, validation and testing.

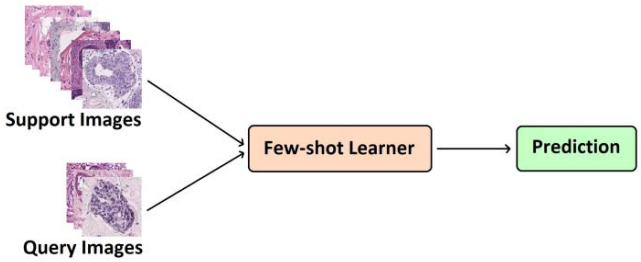

Fig. 2: Block Diagram of Few-shot learning

A. Prototypical Networks

The study employs the BRACS dataset for training, the BACH dataset for validation, BrakeHis for train-validation, and LC25000 data for testing. In the pre-processing step, the study resizes the image to a 28x28 size and performs image rotation at 90, 180, 270, and 360 degrees for data augmentation. Images are also converted from RGB to grayscale. The study uses the ADAM optimizer for optimization. The learning rate is 0.01. The architecture's training epochs are 10000. The model's learning rate starts to decay after twenty epochs. The model runs for 1000 iterations for testing. Fig. 2, Fig. 3 and Fig. 5 display the results of the experiments in 2-way 2-shot 2-query mode, 3-way 3-shot 3-query mode and 5-way 5-shot 5-query mode, respectively. As shown in the figures, it is clear that the model achieves the highest accuracy of 74.575% in the 2-way 2-shot 2 query mode. The model achieves an accuracy of 61.9889% and 45.762% in 3-way 3-shot 3-query mode and 5-way 5-shot 5-query mode, respectively. As the number of classes increases, the accuracy of the model decreases. The training accuracy is also decreasing in every mode with an increase in classes. Fig. 6 and Fig. 7 show the training accuracy and loss of the model, respectively.

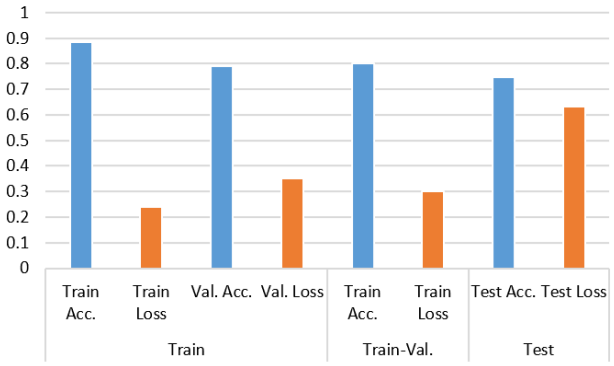

Fig. 3: Prototypical networks in the 2-way 2-shot 2-query mode

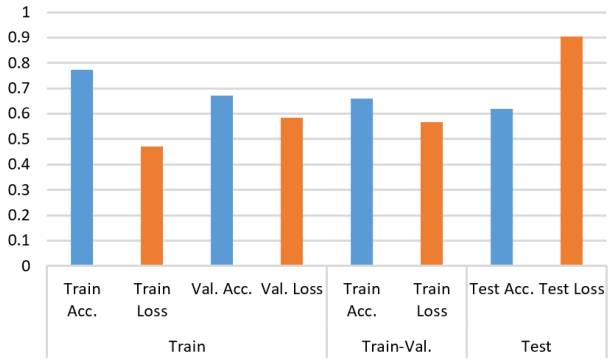

Fig. 4: Prototypical networks in the 3-way 3-shot 3-query mode

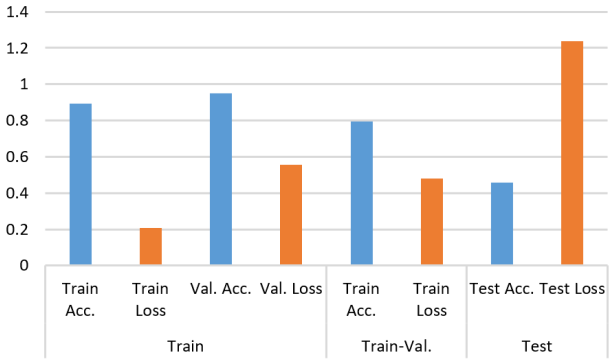

Fig. 5: Prototypical networks in the 5-way 5-shot 5-query mode

B. Model Agnostic Meta Learning

The study employs the BRACS dataset for training, the BACH and BrakeHis datasets for validation, and LC25000 data for testing. The pre-processing is similar to the prototypical networks. The study uses a learning rate of 0.001. The model trains for 15000 epochs. The architecture employs 125 tasks to sample per meta-update. The step size alpha for inner gradient update is 0.001. The number of inner gradient updates in training is 1. The model runs for 15000 iterations for testing.

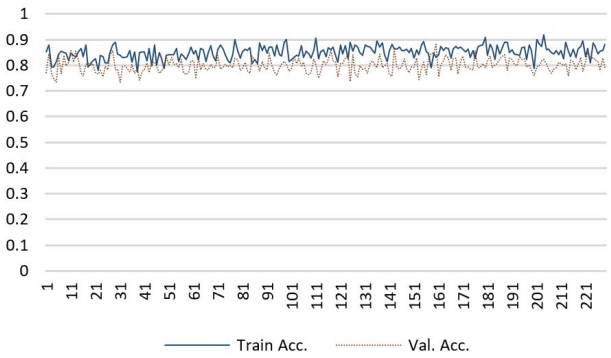

Fig. 6: Training accuracy of Prototypical networks in the 2-way 2-shot 2-query mode

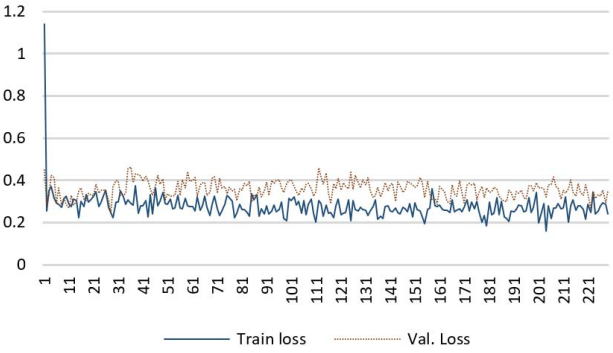

Fig. 7: Training loss of Prototypical networks in the 2-way 2-shot 2-query mode

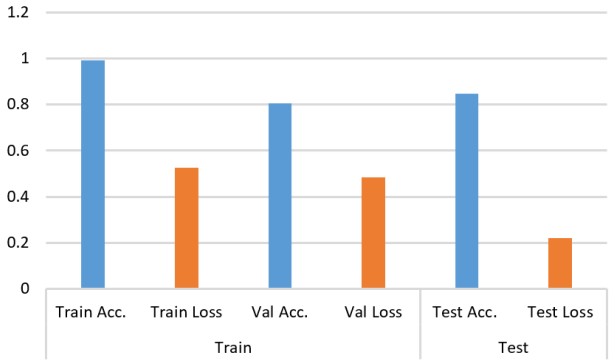

Fig. 8: MAML in the 2-way 2-shot 2-query mode

Fig. 8 displays the results of the experiments in 2-way 2-shot 2-query mode. The model achieves the accuracy of 84.56% in the 2-way 2-shot 2 query mode. Fig. 9 and Fig. 10 show the training accuracy and loss of the model, respectively. Fig.10 displays the result of the experiments in 3-way 3-shot 3-query mode. The model achieves the accuracy of 60.72% in 3-way 3-shot 3-query mode.

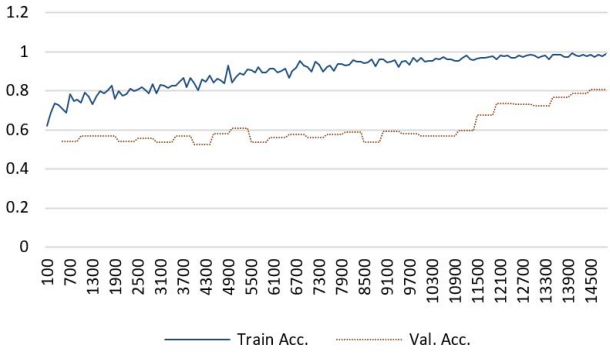

Fig. 9: Training accuracy of MAML in the 2-way 2-shot 2-query mode

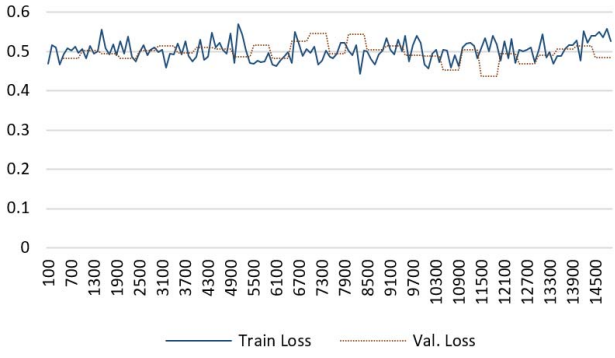

Fig. 10: Training loss of MAML in the 2-way 2-shot 2-query mode

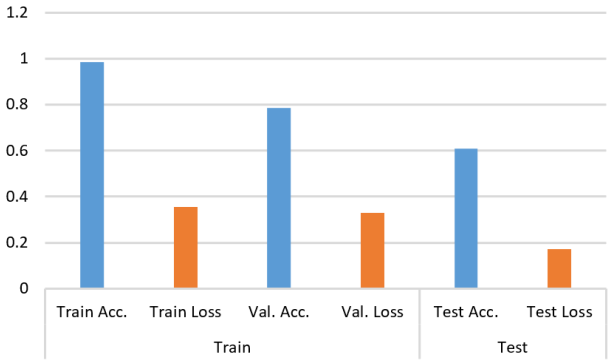

Fig. 11: MAML in the 3-way 3-shot 3-query mode

VI. Conclusion

This study shows the effectiveness of FSL in histopathology image classification. It demonstrates that a model can be trained with a few samples on different datasets. The results show the generalization of the models. A few-shot model learns the features of the dataset and better generalises them to a new dataset. In the experiment, MAML achieved the highest accuracy of 84.56% in the 2-way 2-shot 2 query mode. The study demonstrates the training of FSL models on four different datasets, which is the first time ever in FSL. The outcomes of

the study demonstrate that there is still an opportunity for enhancing performance. Future research will look into the effects of various FSL models on histopathology classification. It is still an assumption that the change in the backbone of CNN may improve the performance. For a more thorough evaluation of the FSL-based approach's potential, researchers may test the algorithms on different disease datasets.

Acknowledgement: This is supported by ISEA Project, Phase-II, MeitY, Govt. of India.

REFERENCES

[1] "Cancer." World Health Organization, WHO, 21 Sept. 2021, www.who.int/news-room/fact-sheets/detail/cancer.

[2] "Cancer." Our World In Data, 22 Nov. 2021, www.ourworldindata.org/cancer.

[3] Lusted, L. B. (1955). Medical electronics. New England Journal of Medicine, 252(14), 580-585.

[4] Lodwick, G. S., Keats, T. E., & Dorst, J. P. (1963). The coding of roentgen images for computer analysis as applied to lung cancer. Radiology, 81(2), 185-200.

[5] Borkowski, A. A., Bui, M. M., Thomas, L. B., Wilson, C. P., DeLand, L. A., & Mastorides, S. M. (2019). Lung and colon cancer histopathological image dataset (lc25000). arXiv preprint arXiv:1912.12142.

[6] Nishio, M., Nishio, M., Jimbo, N., & Nakane, K. (2021). Homology-Based Image Processing for Automatic Classification of Histopathological Images of Lung Tissue. Cancers, 13(6), 1192.

[7] Teh, E. W., & Taylor, G. W. (2020, April). Learning with less data via weakly labeled patch classification in digital pathology. In 2020 IEEE 17th International Symposium on Biomedical Imaging (ISBI) (pp. 471-475). IEEE.

[8] Movshovitz-Attias, Y., Toshev, A., Leung, T. K., Ioffe, S., & Singh, S. (2017). No fuss distance metric learning using proxies. In Proceedings of the IEEE International Conference on Computer Vision (pp. 360-368). ICCV.

[9] Medela, A., Picon, A., Saratxaga, C. L., Belar, O., Cabezón, V., Cicchi, R., & Glover, B. (2019). Few shot learning in histopathological images: reducing the need of labeled data on biological datasets. In 2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019) (pp. 1860-1864). IEEE.

[10] Foucart, A., Debeir, O., & Decaestecker, C. (2018). Artifact identification in digital pathology from weak and noisy supervision with deep residual networks. In 2018 4th International Conference on Cloud Computing Technologies and Applications (Cloudtech) (pp. 1-6). IEEE.

[11] Shaikh, N. N., Wasag, K., & Nie, Y. (2022, March). Artifact Identification in Digital Histopathology Images Using Few-Shot Learning. In 2022 IEEE 19th International Symposium on Biomedical Imaging (ISBI) (pp. 1-4). IEEE.

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

[13] Finn, C., Abbeel, P., & Levine, S. (2017, July). Model-agnostic meta-learning for fast adaptation of deep networks. In International conference on machine learning (pp. 1126-1135). PMLR.

[14] Spanhol, F. A., Oliveira, L. S., Petitjean, C., & Heutte, L. (2015). A dataset for breast cancer histopathological image classification. Ieee transactions on biomedical engineering,, 1455-1462. IEEE.

[15] Brancati, N., Anniciello, A. M., Pati, P., Riccio, D., Scognamiglio, G., Jaume, G., ... & Frucci, M. (2021). Bracs: A dataset for breast carcinoma subtyping in h&e histology images. arXiv preprint.

[16] Aresta, G., Araújo, T., Kwok, S., Chennamsetty, S. S., Safwan, M., Alex, V., & Aguiar, P. (2019). Bach: Grand challenge on breast cancer histology images. Medical image analysis, 122-139.

[17] Das, U. K., Sikder, J., Salma, U., & Anwar, A. S. (2021). Intelligent Cancer Detection System. In 2021 International Conference on Intelligent Technologies (CONIT) (pp. 1-6). IEEE.

[18] Masud, M., Sikder, N., Nahid, A. A., Bairagi, A. K., & AlZain, M. A. (2021). A machine learning approach to diagnosing lung and colon cancer using a deep learning-based classification framework. Sensors, 21(3), 748.

[19] Toğaçar, M. (2021). Disease type detection in lung and colon cancer images using the complement approach of inefficient sets. Computers in Biology and Medicine, 137, 104827.

[20] Ali, M., & Ali, R. (2021). Multi-Input Dual-Stream Capsule Network for Improved Lung and Colon Cancer Classification. Diagnostics, 11(8), 1485.

[21] Garg, S., & Garg, S. (2020). Prediction of lung and colon cancer through analysis of histopathological images by utilizing Pre-trained CNN models with visualization of class activation and saliency maps. In 2020 3rd Artificial Intelligence and Cloud Computing Conference (pp. 38-45).

[22] Phankokkruad, M. (2021). Ensemble Transfer Learning for Lung Cancer Detection. In 2021 4th International Conference on Data Science and Information Technology, (pp. 438- 442).

[23] Lin, J., Han, G., Pan, X., Liu, Z., Chen, H., Li, D., ... & Han, C. (2022). PDBL: Improving Histopathological Tissue Classification With Plug-and-Play Pyramidal Deep-Broad Learning. IEEE Transactions on Medical Imaging, 41(9), 2252-2262.

[24] Mohalder, R. D., & Talukder, K. H. (2021). Deep Learning Based Colorectal Cancer (CRC) Tumors Prediction. In 2021 12th International Conference on Computing Communication and Networking Technologies (ICCCNT) (pp. 01-06). IEEE

[25] Fan, J., Lee, J., & Lee, Y. (2021). A Transfer Learning Architecture Based on a Support Vector Machine for Histopathology Image Classification. Applied Sciences, 11(14), 6380

[26] Adu, K., Yu, Y., Cai, J., Owusu‐Agyemang, K., Twumasi, B. A., & Wang, X. (2021). DHS‐CapsNet: Dual horizontal squash capsule networks for lung and colon cancer classification from whole slide histopathological images. International Journal of Imaging Systems and Technology, 31(4), 2075-2092.