



HAL
open science

A transductive few-shot learning approach for classification of digital histopathological slides from liver cancer

Aymen Sadraoui, Ségolène Martin, Eliott Barbot, Astrid Laurent-Bellue, Jean-Christophe Pesquet, Catherine Guettier, Ismail Ben Ayed

► To cite this version:

Aymen Sadraoui, Ségolène Martin, Eliott Barbot, Astrid Laurent-Bellue, Jean-Christophe Pesquet, et al.. A transductive few-shot learning approach for classification of digital histopathological slides from liver cancer. 2023. hal-04305713v1

HAL Id: hal-04305713

<https://hal.science/hal-04305713v1>

Preprint submitted on 24 Nov 2023 (v1), last revised 9 Mar 2024 (v2)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

A TRANSDUCTIVE FEW-SHOT LEARNING APPROACH FOR CLASSIFICATION OF DIGITAL HISTOPATHOLOGICAL SLIDES FROM LIVER CANCER

Aymen Sadraoui^{1*} Ségolène Martin^{1*} Elliott Barbot¹ Astrid Laurent-Bellue^{2,3,4}
Jean-Christophe Pesquet¹ Catherine Guettier^{2,3,4} Ismail Ben Ayed⁵

¹ Centre de Vision Numérique, Université Paris-Saclay, Inria, CentraleSupélec, Gif-sur-Yvette, France

²Department of Pathology, AP-HP. Hôpital Bicêtre, Le Kremlin-Bicêtre, France

³Faculté de Médecine, Université Paris-Saclay, Le Kremlin-Bicêtre, France

⁴INSERM U1193, Villejuif, France

⁵ École de Technologie Supérieure de Montréal, Montréal, Canada

ABSTRACT

This paper presents a new approach for classifying 2D histopathology patches using few-shot learning. The method is designed to tackle a significant challenge in histopathology, which is the limited availability of labeled data. By applying a sliding window technique to histopathology slides, we illustrate the practical benefits of transductive learning (i.e., making joint predictions on patches) to achieve consistent and accurate classification. Our approach involves an optimization-based strategy that actively penalizes the prediction of a large number of distinct classes within each window. We conducted experiments on histopathological data to classify tissue classes in digital slides of liver cancer, specifically hepatocellular carcinoma. The initial results show the effectiveness of our method and its potential to enhance the process of automated cancer diagnosis and treatment, all while reducing the time and effort required for expert annotation.

Index Terms— histopathology, digital slides, few-shot

1. INTRODUCTION

In clinical settings, histopathology images are a critical primary source of information for pathologists to perform cancer diagnostics and choose treatment strategies. With the widespread adoption of digital pathology, it has become a standard practice to digitize histology slides into high-resolution images called Whole Slide Images (WSIs). WSIs have initiated a new era offering considerable opportunities for using AI assistance systems [1, 2]. In particular, supervised deep learning methods based on conventional neural networks (CNNs) have made great strides in cancer research [3, 4]. However, the success of classical supervised learning approaches depends on the availability of extensive annotated training data. Unlike natural images, which can be annotated via crowd-sourcing, histopathology necessitates expert

pathologists' accurate annotations of gigapixel-sized images. Due to the time-consuming nature of the labeling process, histopathology datasets tend to be limited in size, which poses significant challenges for training machine learning models [5]. Moreover, WSIs can exhibit variability due to staining techniques, tissue preparation, and image quality [6], affecting supervised model performance. Furthermore, supervised learning models can encounter difficulties when confronted with imbalanced data, a common scenario in histopathology. Non-uniform class distribution may produce biased results and compromise the model performance [7].

Few-shot learning methods address the limitations found in traditional supervised learning techniques, providing efficient models capable of generalizing from a small set of labeled examples. These methods not only prove to be scalable, but also significantly reduce costs and time consumption.

Transductive few-shot learning [8, 9], a particularly appealing category within this field, has a distinct advantage. Unlike supervised classification methods, which often treat each data sample independently, transductive methods make predictions on a set of samples collectively. This is especially useful when dealing with localized regions in medical imaging. It allows us to leverage homogeneity and spatial coherence across multiple patches in such a region to enhance the classification accuracy and reliability.

In this work, we introduce a novel transductive few-shot learning approach for histopathological image classification. To our knowledge, it is the first of this kind in the field [10]. Our main contributions are summarized below.

- We apply a sliding window technique to WSIs, establishing a practical scenario where the advantages of transductive few-shot learning are clearly demonstrated.
- Inspired by previous work [8], we develop an optimization-based method for few-shot classification of histopathological patches.
- We validate our approach by tests on the most frequent liver cancer (i.e., hepatocellular carcinoma, HCC), show-

* Equal contributions

casing its effectiveness and confirming its high potential for practical application.

The paper is organized as follows. Section 2 describes the medical context. In Section 3 we detail the few-shot methodology and the proposed algorithm. Finally, Section 4 is dedicated to the experimental results and Section 5 to the conclusion.

2. MEDICAL MULTICLASS PROBLEM

We used HCC WSIs stained with HES (Hemaloxilin-Eosin-Saffron) and digitized at $40\times$ magnification. More precisely, we aim to classify local tissues into the following five classes:

1. Non-Tumor Liver (NT): Liver sections that are not affected by HCC but may be affected by cirrhosis.
2. Hemorrhagic tissue (RE): a non-tumoral pattern characterized by blood cell suffusion.
3. Tumor tissue with macro-trabecular architecture (AM): An aggressive pejorative tumor type characterized by trabeculae of more than ten cells thick.
4. Tumor tissue with Vessels Encapsulating Tumor Clusters architecture (VE): An aggressive pejorative tumor type characterized by tumor cells arranged in small clusters and surrounded by endothelial cells.
5. Conventional trabecular architecture (AN): A non-pejorative tumoral pattern commonly found in HCC patients.

The distinction between tumor and non-tumor areas and the evaluation of pejorative tumor areas provide insightful information to medical doctors.

In this paper, 28 patients from a previously formed cohort of 108 patients with HCC were selected from usable HES slides from Kremlin-Bicêtre Hospital, France, and manually annotated by two skilled pathologists in the five above categories. Annotated WSIs were then tiled into 1728×1728 patches. Data distribution per class is displayed in Table 1.

Class	NT	RE	AM	VE	AN
Percentage	26%	14%	8%	12%	40%

Table 1. Data distribution per class.

3. PROPOSED METHOD

3.1. Problem formulation

Few-shot methods typically involve a two-step process [11]: first, a neural network, pre-trained on a comprehensive and generic dataset, extracts features from the images of interest. Then, a specifically designed classifier is applied to these extracted features to perform the classification task.

We start by introducing the notation for the few-shot classification challenge at hand. The pre-trained network encoder, denoted by Φ , is crucial for feature extraction. Typically, it

has been trained on a dataset $\mathcal{D}_{\text{base}}$ encompassing a broad spectrum of images, potentially inclusive of various WSIs from a multitude of organs and medical facilities. Still, it may not precisely encapsulate the exact categories of our specific classification tasks.

The few-shot dataset consists of N images spanning across K distinct classes. In our context of few-shot classification for histopathological images, K equals 5. Within the dataset, a subset, referred to as the *support set* with index set $\mathbb{S} \subseteq \{1, \dots, N\}$, encompasses the feature samples $(\mathbf{x}_n)_{n \in \mathbb{S}}$ and their respective one-hot-encoded labels $(\mathbf{y}_n)_{n \in \mathbb{S}}$.¹ The support set is constituted by s -shots (labeled examples) for each class. In contrast, the *query set* with indices in $\mathbb{Q} = \{1, \dots, N\} \setminus \mathbb{S}$, comprises a batch of unlabeled samples $(\mathbf{x}_n)_{n \in \mathbb{Q}}$.

The goal is to accurately predict the labels for the samples of the query set under the supervision of the support set. To achieve this, the representations $(\mathbf{z}_n = \Phi(\mathbf{x}_n))_{1 \leq n \leq N}$ generated by the feature extractor are fed into our few-shot classifier.

3.2. Transductive methodology

One of the primary advantages of few-shot learning methodologies, when contrasted with traditional supervised learning techniques, lies in their ability to collectively infer from an entire batch of $|\mathbb{Q}| > 1$ query instances simultaneously rather than evaluating each instance independently. In the lexicon of few-shot learning, this methodology is referred to as *transductive learning* [12, 13]. Transductive few-shot methods are designed to make joint predictions for the entire batch of query samples within each specific few-shot task. This approach takes full advantage of the statistical properties inherent to the query set of a task, employing shared information across instances to enhance generalization and accuracy. Empirical studies have demonstrated that batch-based inference on unlabeled instances, as opposed to individual sample evaluation, results in substantial improvements in prediction accuracy [14].

In the field of microscopy analysis, where spatial pattern recognition is crucial, transductive few-shot approaches exhibit significant potential. Commonly, in a single WSI, it is observed that architectures belonging to the same class tend to cluster spatially, forming homogeneous regions. To leverage this spatial coherence, our strategy involves selecting a window of dimensions $S \times S$ on the microscope slide, as depicted in Figure 1. Each window comprises overlapping mini-patches, each of dimensions $s \times s$, constituting the query set for our few-shot task. The underlying assumption here is that each window encapsulates a few (typically, one or two) distinct classes, allowing each mini-patch to serve as an additional (unlabeled) instance of these classes. By sliding the

¹For every $n \in \mathbb{S}$ and $k \in \{1, \dots, K\}$, $y_{n,k} = 1$ if \mathbf{x}_n is an instance of class k , and $y_{n,k} = 0$ otherwise.

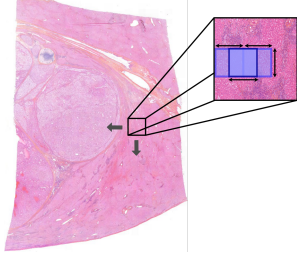


Fig. 1. Scanning of the slide with a sliding window.

window across the entire WSI, we facilitate comprehensive predictions across its entirety.

3.3. Minimization problem

Our method estimates the optimal class assignments of each small patch within the window while limiting the number of predicted classes, thereby acknowledging the spatial coherence in such samples.

By refining the method presented in [8], we approach the few-shot classification challenge through a minimization problem, seeking optimal solutions for the one-hot-encoded assignments $\mathbf{U} = (\mathbf{u}_n)_{1 \leq n \leq |\mathbb{Q}|} \in (\Delta_K)^{|\mathbb{Q}|}$ and the class centroids $\mathbf{W} = (\mathbf{w}_k)_{1 \leq k \leq K} \in (\mathbb{R}^d)^K$, where Δ_K represents the unit simplex set in \mathbb{R}^K . The problem is mathematically formulated as

$$\begin{aligned} & \underset{\mathbf{U}, \mathbf{W}}{\text{minimize}} && f(\mathbf{U}, \mathbf{W}) + g(\mathbf{U}) + \lambda h(\mathbf{U}), && (1) \\ & \text{subject to} && (\forall n \in \mathbb{Q}) \quad \mathbf{u}_n \in \Delta_K, \\ & && (\forall n \in \mathbb{S}) \quad \mathbf{u}_n = \mathbf{y}_n, \end{aligned}$$

with λ is a positive regularization parameter. Here, f represents the data-fidelity term, reflecting the assumption that the data follows a multivariate Gaussian distribution and integrating supervision from the support set. Formally, we define

$$\begin{aligned} f(\mathbf{U}, \mathbf{W}) = & \frac{1}{2} \sum_{k=1}^K \sum_{n=1}^N u_{n,k} (\mathbf{w}_k - \mathbf{z}_n)^\top \hat{\mathbf{S}}_k (\mathbf{w}_k - \mathbf{z}_n) \\ & - \frac{1}{2} \sum_{k=1}^K \sum_{n=1}^N u_{n,k} \ln \det(\hat{\mathbf{S}}_k) \end{aligned} \quad (2)$$

where, for every $k \in \{1, \dots, K\}$, $\hat{\mathbf{S}}_k$ is a symmetric positive matrix corresponding to a sparse approximation of inverse of the empirical covariance matrix of class k , computed from the support set with a Graphical Lasso approach [15]. In addition, g represents an entropic barrier on the assignments, facilitating closed-form updates in the forthcoming algorithm. It is expressed as

$$g(\mathbf{U}) = \sum_{k=1}^K \sum_{n \in \mathbb{Q}} u_{n,k} \ln u_{n,k}. \quad (3)$$

Finally, the penalty function h is central to our approach: it acts as a partition complexity term, encouraging a minimal number of classes to be predicted within the window:

$$h(\mathbf{U}) = - \sum_{k=1}^K \pi_k \ln(\pi_k), \quad (4)$$

where, for every $k \in \{1, \dots, K\}$, $\pi_k = \frac{1}{|\mathbb{Q}|} \sum_{n \in \mathbb{Q}} u_{n,k}$ denotes the proportion of samples of class k in the query set.

3.4. Algorithm

To address the minimization problem outlined in Equation (1), we propose an algorithm that alternates minimization steps with respect to the variables \mathbf{U} and \mathbf{W} . Our iterative approach, detailed in Algorithm 1, shares similarities with the technique presented in [8], the primary distinction being the introduction of inverse covariance matrices. Given these similarities, we direct the reader to [8] for more details on our methodology and the convergence guarantees of the algorithm.

Algorithm 1: PADDLE-Cov

Initialize $\mathbf{W}^{(0)}$ as the means computed on the support set and for all $k \in \{1, \dots, K\}$,

$$\pi_k^{(0)} = \frac{1}{|\mathbb{Q}|} \sum_{n \in \mathbb{Q}} u_{n,k}^{(0)}$$

for $\ell = 1, 2, \dots$, do

$$\begin{aligned} \mathbf{u}_n^{(\ell)} = & \text{softmax} \left(\left(-\frac{1}{2} (\mathbf{w}_k - \mathbf{z}_n)^\top \hat{\mathbf{S}}_k (\mathbf{w}_k - \mathbf{z}_n) \right. \right. \\ & \left. \left. + \frac{1}{2} \ln \det(\hat{\mathbf{S}}_k) + \frac{\lambda}{|\mathbb{Q}|} \ln \pi_k^{(\ell)} \right)_k \right), \quad \forall n \in \mathbb{Q}, \\ \mathbf{w}_k^{(\ell+1)} = & \frac{\sum_{n=1}^N u_{n,k}^{(\ell+1)} \mathbf{z}_n}{\sum_{n=1}^N u_{n,k}^{(\ell+1)}}, \quad \forall k \in \{1, \dots, K\}, \\ \pi_k^{(\ell+1)} = & \frac{1}{|\mathbb{Q}|} \sum_{n \in \mathbb{Q}} u_{n,k}^{(\ell+1)}, \quad \forall k \in \{1, \dots, K\}. \end{aligned}$$

4. EXPERIMENTS

4.1. Experimental setting

In our experimental setup, we leverage the pre-trained model from [16], trained on diverse histopathological images. We structure our few-shot tasks using a sliding window of dimensions 5184×5184 , containing mini-patches of size 1728×1728 downsampled to a resolution of 512×512 . This results in query sets of 25 samples each. The support set comprises the annotated patches of the 28 train patients, and we set the penalty parameter λ to 1250 using validation slides. Preprocessing includes Reinhard color normalization to mitigate staining variability [17].

	Accuracy (%)	F1-score (%)
SimpleShot [18]	48.9	46.4
Baseline [11]	74.4	72.0
α -TIM [9]	56.0	56.9
PADDLE [8]	51.0	48.9
PADDLE-Cov ($\lambda = 0$)	77.3	73.8
PADDLE-Cov	79.3	75.5

Table 2. Evaluation of our approach against other few-shot methods for histopathological patch classification regarding accuracy and F1-score.

4.2. Results

4.2.1. Validation on annotated test data

Our initial evaluation focuses on the entire collection of labeled patches, which we refer to as windows, from the test set slides from 13 patients. Notably, each window is exclusively composed of mini-patches associated with a single class. In this context, we benchmark our approach against two *inductive* few-shot methodologies, SimpleShot [18] and Baseline [11], which conduct inference on each mini-patch independently, as well as the state-of-the-art transductive method α -TIM [9]. In addition we provide an ablation of the terms in our classifying objective (1), evaluating the original PADDLE method (with identity covariances) and the PADDLE-Cov for $\lambda = 0$. The outcomes of this comparative analysis are given in Table 2. Our method surpasses the other approaches, highlighting the benefits of using an appropriate Gaussian metric and of transductive inference.

4.2.2. Inference on a Whole Slide Image (WSI)

In our second evaluation, we aim to compare the predictions made by our 5-class few-shot classifier trained on 28 patients with those of a 3-class fully supervised model on WSIs. The 3-class fully supervised model is a CNN based on ResNet34, which was trained using 800K patches, based on the 87 patients cohort, to classify tissues as non-tumor (NT), non-pejorative tumor (AN), or pejorative tumor (VE+AM). Creating a 5-class supervised model was hardly achievable, as two classes are notably under-represented. Figures 2 and 3 display the predictions of both models on a WSI where the colored squares represent the annotations (ground truth) made by the pathologists.

Both models reliably identify non-tumoral (green squares) and pejorative regions (orange/brown squares), while the conventional trabecular architecture (yellow squares) is better detected by the few-shot model on the WSI in Figure 2. Moreover, training the few-shot model on 5 classes enables detailed detection of the architectures, which the 3-class model can not achieve. In particular, our model accurately distinguishes the VE architecture within the pejorative regions, providing 100% certainty in differentiating VE from

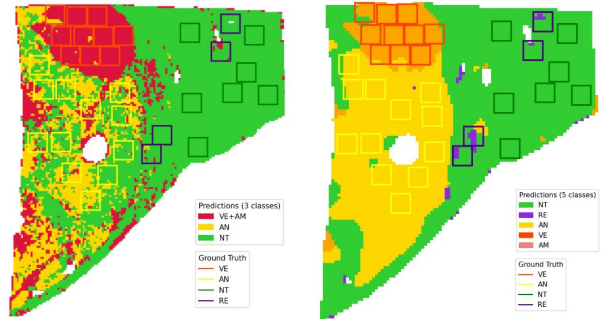


Fig. 2. (left) Predictions made by the 3-class fully supervised model. (right) Predictions made by the few-shot 5-class model.

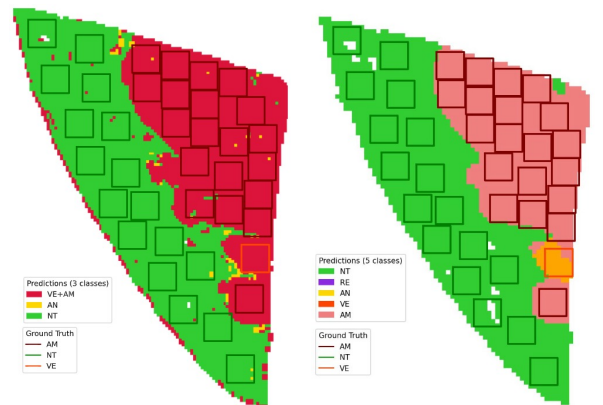


Fig. 3. (left) Predictions made by the 3-class fully supervised model. (right) Predictions made by the few-shot 5-class model.

AM in both WSIs. Additionally, it detected hemorrhagic regions (RE, purple squares), which were logically misclassified by the fully supervised model in Figure 2. Lastly, the 5-class model exhibits remarkable proficiency in defining homogeneous regions across the entire WSIs, unlike the 3-class model, which analyzes individual patches independently. The 5-class model contextual understanding allows for consideration of interdependencies between neighboring patches, leading to a more cohesive interpretation of the data.

5. CONCLUSION

To wrap up, we have introduced an innovative transductive few-shot learning method tailored to classify histopathological images. This approach effectively overcomes significant obstacles, notably data scarcity and class imbalance. Our study emphasizes the adaptability and promise of our method in the domain of biomedical imaging. Its success not only emphasizes the feasibility of our approach in tackling practical challenges but also paves the way for its wider application in various medical imaging scenarios.

6. ETHICS APPROVAL

The study conformed to the General Data Protection Regulation (GDPR) and was approved by the Institutional Review Board of Mondor Hospital (IRB#00011558) (notification number: 2022-135).

7. ACKNOWLEDGMENTS

We would like to thank Dr Laura Claude of the Department of Pathology of the CHU de Rouen, the surgical team of the Centre Hépatobiliaire of the Hôpital Paul Brousse, and the technicians of the Department of Pathology of the Hôpital Bicêtre.

8. REFERENCES

- [1] Neofytos Dimitriou, Ognjen Arandjelovic, and Peter D. Caie, “Deep learning for Whole Slide Image analysis: An overview,” *Clinical Orthopaedics and Related Research*, vol. abs/1910.11097, 2019.
- [2] Kailai Xiang, Baihui Jiang, and Dong Shang, “The overview of the deep learning integrated into the medical imaging of liver: A review,” *Hepatology International*, vol. 15, pp. 868–880, 2021.
- [3] Hoo-Chang Shin, Holger R Roth, Mingchen Gao, Le Lu, Ziyue Xu, Isabella Nogues, Jianhua Yao, Daniel Mollura, and Ronald M Summers, “Deep convolutional neural networks for computer-aided detection: CNN architectures, dataset characteristics and transfer learning,” *IEEE Transactions on Medical Imaging*, vol. 35, no. 5, pp. 1285–1298, 2016.
- [4] Charlie Saillard, Benoit Schmauch, Oumeima Laifa, Matahi Moarii, Sylvain Toldo, Mikhail Zaslavskiy, Elodie Pronier, Alexis Laurent, Giuliana Amaddeo, H el ene Regnault, et al., “Predicting survival after hepatocellular carcinoma resection using deep learning on histological slides,” *Hepatology*, vol. 72, no. 6, pp. 2000–2013, 2020.
- [5] Michael Cooper, Zongliang Ji, and Rahul G Krishnan, “Machine learning in computational histopathology: Challenges and opportunities,” *Genes, Chromosomes and Cancer*, 2023.
- [6] Martin J Van Den Bent, “Interobserver variation of the histopathological diagnosis in clinical trials on glioma: a clinician’s perspective,” *Acta Neuropathologica*, vol. 120, no. 3, pp. 297–304, 2010.
- [7] Justin M Johnson and Taghi M Khoshgoftaar, “Survey on deep learning with class imbalance,” *Journal of Big Data*, vol. 6, no. 1, pp. 1–54, 2019.
- [8] S egol ene Martin, Malik Boudiaf, Emilie Chouzenoux, Jean-Christophe Pesquet, and Ismail Ayed, “Towards practical few-shot query sets: Transductive minimum description length inference,” *Advances in Neural Information Processing Systems*, vol. 35, pp. 34677–34688, 2022.
- [9] Olivier Veilleux, Malik Boudiaf, Pablo Piantanida, and Ismail Ben Ayed, “Realistic evaluation of transductive few-shot learning,” *Advances in Neural Information Processing Systems*, vol. 34, 2021.
- [10] Joanna Szolomicka and Urszula Markowska-Kaczmar, “An overview of few-shot learning methods in analysis of histopathological images,” *Advances in Smart Healthcare Paradigms and Applications*, pp. 87–113, 2023.
- [11] Wei-Yu Chen, Yen-Cheng Liu, Zsolt Kira, Yu-Chiang Frank Wang, and Jia-Bin Huang, “A closer look at few-shot classification,” in *International Conference on Learning Representations*, 2019.
- [12] John Bronskill, Jonathan Gordon, James Requeima, Sebastian Nowozin, and Richard Turner, “Tasknorm: Rethinking batch normalization for meta-learning,” in *International Conference on Machine Learning*. PMLR, 2020, pp. 1153–1164.
- [13] Yuqing Hu, Vincent Gripon, and St ephane Pateux, “Leveraging the feature distribution in transfer-based few-shot learning,” in *International Conference on Artificial Neural Networks*. Springer, 2021, pp. 487–499.
- [14] Thorsten Joachims, “Transductive inference for text classification using support vector machines,” in *International Conference on Machine Learning*, 1999, vol. 99, pp. 200–209.
- [15] Jerome Friedman, Trevor Hastie, and Robert Tibshirani, “Sparse inverse covariance estimation with the graphical lasso,” *Biostatistics*, vol. 9, no. 3, pp. 432–441, 2008.
- [16] Ozan Ciga, Tony Xu, and Anne Louise Martel, “Self supervised contrastive learning for digital histopathology,” *Machine Learning with Applications*, vol. 7, pp. 100198, 2022.
- [17] E. Reinhard, M. Adhikhmin, B. Gooch, and P. Shirley, “Color transfer between images,” *IEEE Computer Graphics and Applications*, vol. 21, no. 5, pp. 34–41, 2001.
- [18] Yan Wang, Wei-Lun Chao, Kilian Q. Weinberger, and Laurens Van der Maaten, “Simpleshot: Revisiting nearest-neighbor classification for few-shot learning,” *Computer Vision and Pattern Recognition*, 2019.