D2C: Data-Efficient Skin Lesion Classification for Clinical Images by Representation Alignment Using Dermoscopic Resources

Zichen Zhang[1], Naweed Shifai[2], and Oleh Dzyubachyk[1]

[1] Division of Image Processing, Department of Radiology, Leiden University Medical Center, Leiden, the Netherlands
[2] Department of Dermatology, Leiden University Medical Center, Leiden, the Netherlands
o.dzyubachyk@lumc.nl

Abstract. Skin lesion classification is a crucial clinical task in skin image analysis. Dermoscopy is widely adopted for the diagnosis of skin cancer due to its high image quality and the ability to reveal surface structures. However, dermoscopes require a certain level of expertise to use and thus become less effective when experienced dermatologists are scarce. In comparison, clinical images can be acquired with ordinary cameras, which are more accessible and are therefore increasingly utilized. Nevertheless, clinical images are less diagnostic-effective due to their lower image quality and greater imaging variability. In recent years, although public dermoscopic datasets and deep learning models have been successfully established, little research has been conducted on adapting these resources to clinical images. In this paper, we formulate this research topic as a semi-supervised domain adaptation problem and propose a data-efficient framework for skin lesion classification for clinical images. Abundant diagnostic features extracted from dermoscopic images will be established as informative knowledge vaults to facilitate the transfer of diagnostic capability from the dermoscopic image domain to the clinical image domain. Cluster-level and instance-level adaptation strategies are proposed for this knowledge transfer under different granularities. Comprehensive experiments across diverse datasets and settings demonstrate the superiority of our framework in skin lesion classification and data efficiency, showing its potential for clinical applications.

Keywords: Skin lesion classification · Domain adaptation · Representation alignment.

1 Introduction

In recent years, the number of deaths caused by skin cancer has been increasing consistently, which has drawn great attention all over the world [5]. Dermoscopy, a crucial non-invasive imaging modality for skin cancer, plays a vital role in its diagnostic procedures. Dermoscopy usually provides a detailed, zoomed-in, and

surface-reflection-reduced view of the skin lesion, revealing clear surface structures that can significantly aid lesion diagnosis. Conventionally, such a diagnosis is carried out by clinicians using predefined criteria such as the 7-point checklist [9] for melanoma. With the development of deep learning techniques, these data-driven approaches have been applied to automate skin lesion diagnosis, with a primary focus on dermoscopic image classification [14,15,22]. In the past few years, the International Skin Imaging Collaboration (ISIC) has collected large public skin image datasets [3,2,20] and greatly advanced this field.

However, using dermoscopes requires some expertise, and specifically, general practitioners are often not extensively trained on diagnosing skin cancer with dermoscopes. On the contrary, another imaging modality, clinical images captured with ordinary digital cameras, are increasingly utilized because of their simplicity, cost-effectiveness, and easy accessibility. These clinical images provide a broader view of the skin area and can complement dermoscopic images for diagnosis. Nonetheless, deep learning for such clinical skin images is a relatively underexplored area, as much of the attention is drawn to dermoscopic image analysis, and the public dermoscopic datasets are more abundant and organized. Under such circumstances, it is clinically significant to discover how to transfer the dermoscopic resources for clinical image analysis.

We formulate this research question in a semi-supervised domain adaptation (SSDA) setting, where the source domain contains abundant labeled dermoscopic images, and the target domain consists of much fewer labeled clinical images and a relatively larger portion of unlabeled clinical images. The technical objective is to train a deep classifier on these data to achieve decent performance on clinical images with an emphasis on domain adaptation, while the overall clinical objective is to differentiate skin lesions using clinical images with the aid of dermoscopic ones. Various deep learning-based methods have been proposed for SSDA. ECACL [13] proposes two approaches to align two domains. One is to use target-domain prototypes to classify source-domain images, the other is to group features of the same class closer while keeping distances to distinct classes. ProML [11] focuses on pseudo-label refinement by optimal transport to obtain more reliable pseudo labels. It also leverages prototypes with momentum updates to align two domains. JCLS [21] defines sensitivities to improve contrastive learning in key-sample mining. MuVo [10] maintains a feature bank that explicitly stores more relevant features to improve domain adaptation by prototypical learning. SERL [12] proposes to apply mixup regularization to hard samples to enhance the separability of these samples that are near the decision boundaries. As dermoscopic datasets usually contain more data than clinical datasets, conventional purely even-batch-based computation does not fully exploit the abundant information from dermoscopic resources. Besides, the above approaches focus almost entirely on the feature space, neglecting how pixel-space variations across the two domains affect representation alignment, which is important for dermoscopic and clinical images, as their imaging variations exhibit patterns describable with prior knowledge.

In this paper, we propose a deep learning-based SSDA framework to achieve data-efficient skin lesion classification for clinical images. We first train a classifier using labeled dermoscopic images. Next, class-aware dermoscopic feature vaults are built by feeding the dermoscopic images to the classifier and acquiring the dermoscopic features. Then, the trained deep classifier will be adapted to clinical images via an SSDA scheme using data from both domains and the built dermoscopic feature vaults. Several representation alignment strategies are proposed to align the feature spaces of the two domains, especially with the aid of the feature vaults, which provide more comprehensive information than batch-level data, thereby boosting the performance of the classifier on the clinical domain while improving data efficiency. Comprehensive experiments conducted on public datasets with various settings demonstrate that our framework surpasses prior methods in data-efficient skin lesion classification for clinical images.

2 Methods

An overview of our framework is illustrated in Fig. 1, consisting of two stages shown in the upper and lower parts in Fig. 1, respectively. The first stage is to train a deep classifier using a labeled dermoscopic dataset. After that, the dermoscopic images are fed into the classifier to produce dermoscopic features, which are collected into several feature vaults. At the second stage, the trained classifier is adapted to the target clinical domain via the proposed SSDA scheme.

2.1 Dermoscopic Domain Training

The classifier consists of a backbone and two linear layers. The backbone and the first linear layer act as the feature extractor $\mathcal{E}$. In this paper, the image features are the vectors produced by $\mathcal{E}$. These features are sent to the second linear layer, denoted as the classification head $\mathcal{C}$, to obtain unnormalized classification logits. The dermoscopic domain training is performed in a supervised manner, with the employed loss function being cross-entropy. After training, the dermoscopic images from the training set are fed into $\mathcal{E}$ to produce the dermoscopic features. These features are separated by their corresponding image labels and stored separately to serve as learning guidance for the following SSDA scheme. If there are $c \in \mathbb{N}^+$ classes, then there will be c dermoscopic feature vaults.

2.2 Semi-Supervised Domain Adaptation

In this stage, the classifier trained on the source dermoscopic domain will be adapted to the target clinical domain using an SSDA scheme, leveraging data from both domains and the gathered dermoscopic feature vaults. Various representation alignment strategies are proposed to align the feature spaces of the two domains, encouraging domain-invariance and feature compactification.

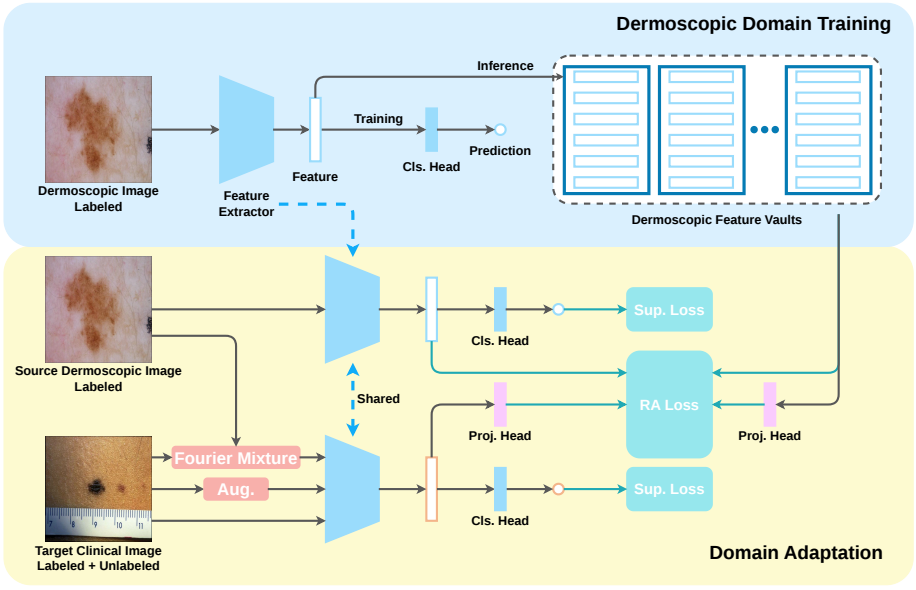

Fig. 1. An overview of our framework. A classifier is first trained solely using dermoscopic resources to produce high-quality dermoscopic diagnostic feature vaults. Next, an SSDA architecture with several domain alignment strategies is introduced to adapt the trained classifier to the clinical image domain.

Supervision from Labels. Since there are image labels for both the source and target domains, cross-entropy is employed for supervised learning on both domains, yielding the following loss function:

$$\mathcal{L}_{sup} = \mathcal{L}_{CE}(\mathcal{C}(\mathcal{E}(\mathbf{X}_s)), y_s) + \mathcal{L}_{CE}(\mathcal{C}(\mathcal{E}(\mathbf{X}_t)), y_t) \tag{1}$$

where $\mathbf{X}_s, \mathbf{X}_t \in \mathbb{R}^{3\times h\times w}$ are images from both domains and $y_s, y_t \in \mathbb{C}$ are the respective image labels. Here, $\mathbb{C} = \{1, ..., c\}$, where c is the number of classes.

Class-Aware Representation Alignment. We first compute the centroids of the dermoscopic feature vaults to obtain the source feature centroids of each class $\mathbf{c}_s^k \in \mathbb{R}^n$, where $k \in \mathbb{C}$ and $n \in \mathbb{N}^+$ is the dimensionality of the feature vectors. At each training iteration, a batch of target-domain images is fed into $\mathcal{E}$ to acquire the target-domain features. We compute their class-wise feature centroids as well, yielding $\mathbf{c}_t^k \in \mathbb{R}^n$. Then, we perform class-aware representation alignment between the two domains with the following loss function:

$$\mathcal{L}_{cara} = \frac{1}{c}\sum_{k=1}^{c} \left\|\mathbf{c}_s^k - \mathbf{c}_t^k\right\|_2^2 \tag{2}$$

where the two domains are more accurately aligned, as $\mathbf{c}_s^k$ are computed from the large feature vaults rather than a batch of data and are therefore less biased.

Instance-Wise Contrastive Alignment. The above $\mathcal{L}_{cara}$ aligns the two domains cluster-wise, and we will also perform instance-wise alignment for fine-grained cross-domain interactions. Instead of applying contrastive learning to the feature space that can affect its structure, a shallow projection head $\mathcal{P}$ is introduced to map the feature space to a closely neighboring space, such that the influence of contrastive learning can be passed to the feature space in a milder way. Given a target-domain feature $\mathbf{f}_t^k \in \mathbb{R}^n$ belonging to class k, the respective positive keys are the source-domain features with the same class, and the negative keys are the ones from both domains with non-k classes. The InfoNCE [16] loss is leveraged to perform such contrastive learning:

$$\mathcal{L}_{iwca} = -\sum_{\mathbf{f}_t^k} log \frac{exp(\mathcal{S}(\mathbf{f}_t^k, \mathbf{f}_s^k)/\tau)}{exp(\mathcal{S}(\mathbf{f}_t^k, \mathbf{f}_s^k)/\tau) + \sum_{i=s,t}\sum_{k' \in \mathbb{C}\setminus k} exp(\mathcal{S}(\mathbf{f}_t^k, \mathbf{f}_i^{k'})/\tau)} \tag{3}$$

where $\mathcal{S}$ is a similarity mapping that has the following formulation:

$$\mathcal{S}(\mathbf{f}_1, \mathbf{f}_2) = \frac{\mathcal{P}(\mathbf{f}_1)}{\|\mathcal{P}(\mathbf{f}_1)\|_2} \cdot \frac{\mathcal{P}(\mathbf{f}_2)}{\|\mathcal{P}(\mathbf{f}_2)\|_2} \tag{4}$$

where $\cdot$ is the vector dot product. Performing contrastive learning in this way maximizes its power, as the number of negative samples is significantly enlarged when using the dermoscopic feature vaults rather than batched data.

Cross-Domain Feature Compactification. Dermoscopic and clinical images exhibit imaging variations in aspects such as field of view, lighting conditions, and surface reflections. In this alignment strategy, we employ an image-level mixture to compact the learned features and facilitate domain alignment. Inspired by [23], we first perform Fast Fourier Transform (FFT) to a source image $\mathbf{X}_s$ and target image $\mathbf{X}_t$, either labeled or unlabeled, deriving two amplitude maps $\mathbf{A}_s$, $\mathbf{A}_t$ and two phase maps. Next, the low-frequency spectra $\mathbf{A}_s^*$, $\mathbf{A}_t^*$ of $\mathbf{A}_s$, $\mathbf{A}_t$, which store mostly the non-content components, are linearly blended as follows:

$$\mathbf{A}_{mix}^* = \alpha \mathbf{A}_s^* + (1-\alpha)\mathbf{A}_t^*, \alpha \sim \mathcal{U}(a, b) \tag{5}$$

where α follows a uniform distribution between $a \in (0, 1)$ and $b \in (0, 1)$. $\mathbf{A}_{mix}^*$ then becomes the new low-frequency spectrum of $\mathbf{A}_t$. Inverse FFT is performed using the mixed amplitude map and the original phase map to obtain the mixed target image $\mathbf{X}_{t,mix}$. We encourage the features of $\mathbf{X}_{t,mix}$ and its original counterpart to be consistent with the following loss function:

$$\mathcal{L}_{cdfc} = \|\mathcal{E}(\mathbf{X}_{t,mix}) - sg(\mathcal{E}(\mathbf{X}_t))\|_2^2 \tag{6}$$

where sg is the stop-gradient operation. In contrast to [23], which uses spectrum replacement, our linear blending method provides a series of mixed images for a smoother alignment process. In this way, the imaging variations across the two domains are effectively interacted to compactify the learned features.

Feature Compactification for Unlabeled Images. Since we have access to unlabeled clinical images, we propose a self-distillation-like alignment strategy for these data. Instead of leveraging the conventional pseudo-labeling in semi-supervised learning, which can easily generate false pseudo-labels, we directly encourage the features extracted from augmented unlabeled images to match those from the original images, compactifying the learned features and making them photometrically and geometrically invariant. This approach aligns with the properties of clinical images, which are captured with mobile digital cameras and exhibit large variation due to camera poses and lighting conditions. The loss function is formulated as follows:

$$\mathcal{L}_{fcui} = \|\mathcal{E}(\mathbf{X}_{t,aug}) - sg(\mathcal{E}(\mathbf{X}_t))\|_2^2 \tag{7}$$

where $\mathbf{X}_{t,aug}$ is an augmented unlabeled clinical image.

Total Loss Function. Additionally, we encourage the class-wise feature clusters to spread narrowly to improve class separability by using the regularization term:

$$\mathcal{L}_{std} = \sum_{i=s,t} \sum_{k \in \mathbb{C}} std(\mathcal{E}(\mathbf{X}_i^k)) \tag{8}$$

where *std* is the standard deviation. The representation alignment loss is:

$$\mathcal{L}_{ra} = \lambda_1 \mathcal{L}_{cara} + \lambda_2 \mathcal{L}_{iwca} + \lambda_3 \mathcal{L}_{cdfc} + \lambda_4 \mathcal{L}_{fcui} + \lambda_5 \mathcal{L}_{std} \tag{9}$$

where λ_1, λ_2, λ_3, λ_4, and λ_5 are balancing factors. The overall loss function is:

$$\mathcal{L} = \mathcal{L}_{sup} + \mathcal{L}_{ra} \tag{10}$$

3 Experiments and Results

3.1 Experimental Settings

Datasets. For the source dermoscopic domain, we utilize the public ISIC-2017 [3] and ISIC-2018 [2,20] datasets. In this paper, we primarily focus on melanoma classification and thus select images from the original datasets that are melanoma or nevus. Therefore, c is 2 and $\mathbb{C} = \{1, 2\}$. After the selection, we acquire 2064 images from ISIC-2017 and 7962 images from ISIC-2018 to train the deep classifier. For the target clinical domain, we collect melanoma and nevus images from the public PAD-UFES-20 [17] and Fitzpatrick17k [7,6] datasets and merge them into a single clinical dataset with 1120 images to enhance data quantity and experimental effectiveness. In this semi-supervised domain adaptation setting, the source domain is either ISIC-2017 or ISIC-2018, and the target domain is the merged clinical dataset. We set two experimental settings for clinical data, using either 5% as labeled data and 55% as unlabeled data or 10% as labeled data and 50% as unlabeled data. The models are tested on the other data. All the images are resized to 512 × 512.

Implementation Details. The backbone employed in this paper is ResNet-18 [8] pretrained on ImageNet [4]. The feature dimension n is 128. The projection head consists of a ReLU layer and two linear layers, with a hidden dimension of 128. The low-frequency spectrum is the central 2×2 area in the shifted spectrum. Factors a, b, λ_1, λ_2, λ_3, λ_4, and λ_5, selected based on typical values empirically, are 0.3, 0.7, 0.1, 0.01, 0.01, 1, and 0.001. The batch size, learning rate, and number of training epochs for the two training stages are 16, 0.001, and 200. We utilize the SGD optimizer with a momentum of 0.9 and a weight decay of 0.0001. The learning rate is tuned with a reduce-on-plateau scheduler. All network training is performed on an NVIDIA GeForce RTX 4090 GPU.

Competing Methods and Evaluation. Five SSDA methods, ECACL [13], ProML [11], JCLS [21], MuVo [10], and SERL [12], are included for comparison. ECACL [13] has two variants, both of which are included. In addition, we add two semi-supervised (SS) approaches, ReCo [24] and FCAT [18], and two supervised domain adaptation (SDA) methods, DivAug [1] and SAT-DA [19], to the experiments to enhance comprehensiveness. SS methods have access to dermoscopic images and use them through supervised learning. For a fair comparison, these methods are reproduced and switched to a uniform ResNet-18 backbone. We utilize AUROC as the evaluation metric.

3.2 Experimental Results

The performances of the methods are listed in Table 1 and Table 2. Our framework achieves the best overall classification performance in the four experimental settings. It outperforms other approaches more in the ISIC-2018 [2,20] adaptation setting, suggesting that the effectiveness of our approach, especially the strategies that use the dermoscopic feature vaults, continues to increase as the source domain data scales up.

Table 1. The performances of adaptation from ISIC-2017 [3] to clinical images.

Methods	Method Types	5% Labeled Data	10% Labeled Data
ECACL-P [13]	SSDA	73.62 ± 1.01	77.57 ± 2.07
ECACL-T [13]	SSDA	76.56 ± 2.89	78.00 ± 2.96
ProML [11]	SSDA	74.56 ± 3.11	78.99 ± 2.73
JCLS [21]	SSDA	76.27 ± 2.67	76.64 ± 3.43
MuVo [10]	SSDA	76.94 ± 2.48	78.45 ± 3.39
SERL [12]	SSDA	73.50 ± 4.20	77.05 ± 1.93
ReCo [24]	SS	74.74 ± 3.61	78.44 ± 2.36
FCAT [18]	SS	75.83 ± 2.57	78.71 ± 3.57
DivAug [1]	SDA	73.58 ± 2.55	76.09 ± 4.51
SAT-DA [19]	SDA	72.35 ± 1.21	75.86 ± 2.97
Ours	SSDA	**78.08 ± 1.91**	**81.68 ± 2.53**

Table 2. The performances of adaptation from ISIC-2018 [2,20] to clinical images.

Methods	Method Types	5% Labeled Data	10% Labeled Data
ECACL-P [13]	SSDA	74.05 ± 2.09	75.24 ± 2.81
ECACL-T [13]	SSDA	73.21 ± 4.05	75.37 ± 2.29
ProML [11]	SSDA	74.46 ± 3.71	78.36 ± 5.22
JCLS [21]	SSDA	72.38 ± 2.82	76.41 ± 0.87
MuVo [10]	SSDA	74.18 ± 2.62	77.97 ± 3.49
SERL [12]	SSDA	71.77 ± 2.66	75.42 ± 2.45
ReCo [24]	SS	78.58 ± 3.98	78.68 ± 4.81
FCAT [18]	SS	73.53 ± 2.67	78.24 ± 2.55
DivAug [1]	SDA	72.91 ± 2.71	76.52 ± 2.20
SAT-DA [19]	SDA	73.80 ± 0.66	74.88 ± 2.84
Ours	SSDA	**80.72 ± 4.10**	**81.98 ± 2.75**

3.3 Ablation Studies

We also conducted ablation studies to verify the effectiveness of each component of our framework, as shown in Table 3. The result in the first row is obtained by applying only supervised learning using the labeled data. From the ablation results, each component removal results in a performance degradation, demonstrating the effectiveness of our representation alignment strategies for SSDA.

Table 3. The ablation studies on adaptation from ISIC-2018 [2,20] to clinical images.

Configurations	5% Labeled Data
Supervised Training	72.49 ± 5.11
W/O $\mathcal{L}_{cara}$	77.34 ± 5.64
W/O $\mathcal{L}_{iwca}$	77.66 ± 3.36
W/O $\mathcal{L}_{cdfc}$	78.22 ± 4.16
W/O $\mathcal{L}_{fcui}$	79.01 ± 4.07
W/O $\mathcal{L}_{std}$	79.96 ± 3.97
Proposed	**80.72 ± 4.10**

4 Conclusion

In this paper, we propose an SSDA framework for data-efficient skin lesion classification for clinical images. The abundant dermoscopic knowledge is effectively transferred to the target clinical domain by various representation alignment strategies and the built dermoscopic feature vaults. Comprehensive experiments under varied settings have demonstrated that our method outperforms the competing approaches. Besides, as all methods require only one forward pass except DivAug [1], which requires two forward passes due to its decoupled design, our method achieves an equivalent or better inference efficiency.

Acknowledgments. This research is funded by a grant from the Hanarth Foundation (Stichting Hanarth Fonds). Naweed Shifai is supported by a personal grant from the Dutch Research Council (NWO) via the research programme Mozaïek 2.0 with file number IIW.1154.24.022 under the grant ID [https://doi.org/10.61686/SIZKF12238].

Disclosure of Interests. The project is supported by a grant from the Hanarth Foundation (Stichting Hanarth Fonds). Naweed Shifai has received a personal grant from the Dutch Research Council (NWO) via the research programme Mozaïek 2.0.

References

1. Chen, Z., Pu, B., Zhao, L., He, J., Liang, P.: Divide and augment: Supervised domain adaptation via sample-wise feature fusion. Information Fusion **115**, 102757 (2025)
2. Codella, N., Rotemberg, V., Tschandl, P., Celebi, M.E., Dusza, S., Gutman, D., Helba, B., Kalloo, A., Liopyris, K., Marchetti, M., et al.: Skin lesion analysis toward melanoma detection 2018: A challenge hosted by the international skin imaging collaboration (isic). arXiv preprint arXiv:1902.03368 (2019)
3. Codella, N.C., Gutman, D., Celebi, M.E., Helba, B., Marchetti, M.A., Dusza, S.W., Kalloo, A., Liopyris, K., Mishra, N., Kittler, H., et al.: Skin lesion analysis toward melanoma detection: A challenge at the 2017 international symposium on biomedical imaging (isbi), hosted by the international skin imaging collaboration (isic). In: 2018 IEEE 15th international symposium on biomedical imaging (ISBI 2018). pp. 168–172. IEEE (2018)
4. Deng, J., Dong, W., Socher, R., Li, L.J., Li, K., Fei-Fei, L.: Imagenet: A large-scale hierarchical image database. In: 2009 IEEE conference on computer vision and pattern recognition. pp. 248–255. Ieee (2009)
5. Garbe, C., Leiter, U.: Melanoma epidemiology and trends. Clinics in dermatology **27**(1), 3–9 (2009)
6. Groh, M., Harris, C., Daneshjou, R., Badri, O., Koochek, A.: Towards transparency in dermatology image datasets with skin tone annotations by experts, crowds, and an algorithm. Proceedings of the ACM on Human-Computer Interaction **6**(CSCW2), 1–26 (2022)
7. Groh, M., Harris, C., Soenksen, L., Lau, F., Han, R., Kim, A., Koochek, A., Badri, O.: Evaluating deep neural networks trained on clinical images in dermatology with the fitzpatrick 17k dataset. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 1820–1828 (2021)
8. He, K., Zhang, X., Ren, S., Sun, J.: Deep residual learning for image recognition. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 770–778 (2016)
9. Herschorn, A.: Dermoscopy for melanoma detection in family practice. Canadian Family Physician **58**(7), 740–745 (2012)
10. Hong, Y., Dong, L., Qiu, X., Xiao, H., Yao, B., Zheng, S., Peng, C.: A multi-view consistency framework with semi-supervised domain adaptation. Engineering Applications of Artificial Intelligence **136**, 108886 (2024)
11. Huang, X., Zhu, C., Chen, W.: Semi-supervised domain adaptation via prototype-based multi-level learning. In: Elkind, E. (ed.) Proceedings of the Thirty-Second International Joint Conference on Artificial Intelligence, IJCAI-23. pp. 884–892. International Joint Conferences on Artificial Intelligence Organization (8

2023). https://doi.org/10.24963/ijcai.2023/98, https://doi.org/10.24963/ijcai.2023/98, main Track
12. Huang, X., Zhu, C., Ren, R., Liu, S., Huang, T.: Source-free semantic regularization learning for semi-supervised domain adaptation. IEEE Transactions on Multimedia **27**, 4227–4239 (2025)
13. Li, K., Liu, C., Zhao, H., Zhang, Y., Fu, Y.: Ecacl: A holistic framework for semi-supervised domain adaptation. In: Proceedings of the IEEE/CVF international conference on computer vision. pp. 8578–8587 (2021)
14. Liu, Z., Xiong, R., Jiang, T.: Clinical-inspired network for skin lesion recognition. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 340–350. Springer (2020)
15. Liu, Z., Xiong, R., Jiang, T.: Multi-level relationship capture network for automated skin lesion recognition. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 153–164. Springer (2021)
16. Oord, A.v.d., Li, Y., Vinyals, O.: Representation learning with contrastive predictive coding. arXiv preprint arXiv:1807.03748 (2018)
17. Pacheco, A.G., Lima, G.R., Salomao, A.S., Krohling, B., Biral, I.P., De Angelo, G.G., Alves Jr, F.C., Esgario, J.G., Simora, A.C., Castro, P.B., et al.: Pad-ufes-20: A skin lesion dataset composed of patient data and clinical images collected from smartphones. Data in brief **32**, 106221 (2020)
18. Shiyan, L., Shuqin, W., Xin, G., Debing, S.: Semi-supervised medical image classification via feature-level multi-scale consistency and adversarial training. Computerized Medical Imaging and Graphics p. 102695 (2025)
19. Sultana, N., Lu, W., Fan, X., Yap, M.H.: Selective alignment transfer for domain adaptation in skin lesion analysis. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 596–606. Springer (2025)
20. Tschandl, P., Rosendahl, C., Kittler, H.: The ham10000 dataset, a large collection of multi-source dermatoscopic images of common pigmented skin lesions. Scientific data **5**(1), 180161 (2018)
21. Tu, K., Wang, Z., Li, J., Zhang, Y.: Semi-supervised domain adaptation via joint contrastive learning with sensitivity. In: Proceedings of the 31st ACM international conference on multimedia. pp. 5645–5654 (2023)
22. Wang, Y., Wang, Y., Cai, J., Lee, T.K., Miao, C., Wang, Z.J.: Ssd-kd: A self-supervised diverse knowledge distillation method for lightweight skin lesion classification using dermoscopic images. Medical Image Analysis **84**, 102693 (2023)
23. Yang, Y., Soatto, S.: Fda: Fourier domain adaptation for semantic segmentation. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. pp. 4085–4095 (2020)
24. Zeng, Q., Lu, Z., Xie, Y., Lu, M., Ma, X., Xia, Y.: Reciprocal collaboration for semi-supervised medical image classification. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 522–532. Springer (2024)