ISIC Skin Image Analysis Workshop
@ CVPR 2020

# Meta-DermDiagnosis: Few-Shot Skin Disease Identification using Meta-Learning

**Kushagra Mahajan**, Monika Sharma, Lovekesh Vig
TCS Research, New Delhi, India

# Introduction

- Disease classification and quick model adaptation in low-data situations and datasets with long-tailed class distributions using meta-learning techniques.
- Few-shot learning techniques such as the gradient based Reptile [1] and distance metric based Prototypical networks [2] are used.
- Evaluated our approach on 3 publicly available skin lesion datasets: ISIC 2018 [3], Derm7pt [4] and SD-198 [5] datasets.
- Obtained significant performance improvement over pre-trained models using meta-learning techniques.
- Incorporated Group Equivariant convolutions (G-convolutions) [6] to improve disease identification as they make the network equivariant to discrete transformations like rotation, reflection and translation.

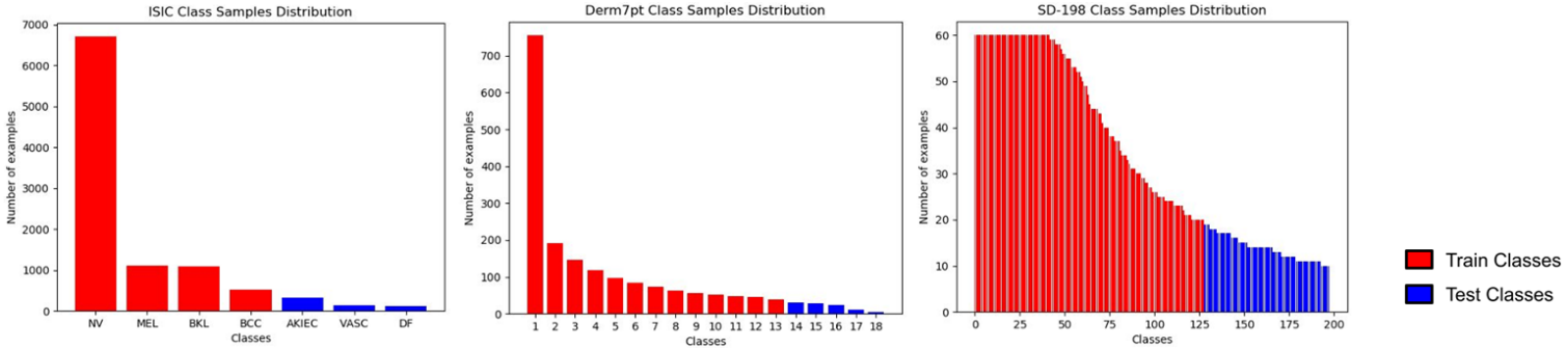
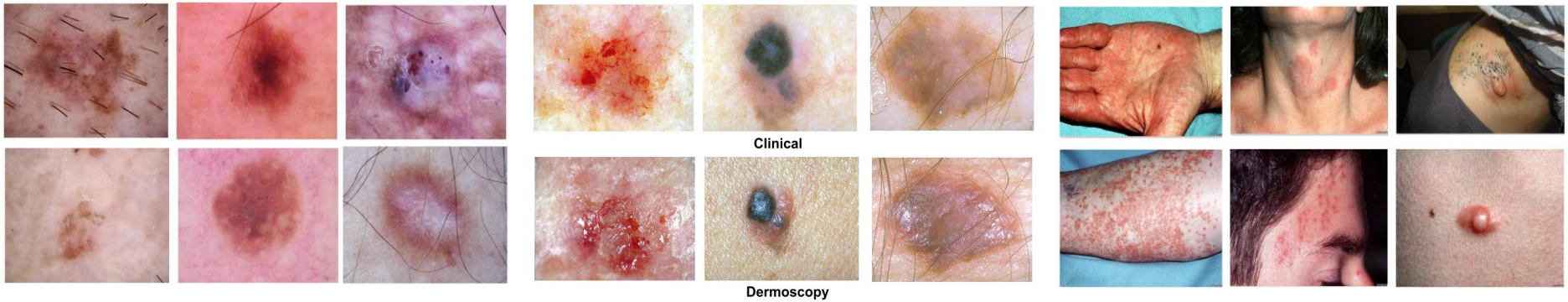



**Figure 1.** Figures showing class distribution in skin lesion datasets: ISIC 2018, Derm7pt and SD-198. The classes towards head of the class distribution (common-diseases), shown in **red**, are taken as train classes and classes at the tail of the distribution (new / rare disease), shown in **blue** color, are chosen as test classes.




**(a) ISIC 2018 [3]** **(b) Derm7pt [4]** **(c) SD-198 [5]**

**Figure 2.** Figure showing some sample images from skin lesion datasets.

# Motivation

- New ailments are continuously being discovered, with lack of sufficient data for accurate classification.
- Annotations of these ailments like skin diseases from images by experienced doctors is very time consuming, labour intensive, costly and error-prone.
- Conventional deep networks tend to fail when there is limited annotated data available since they overfit.
- However, humans can learn quickly from a few examples by leveraging prior knowledge.
- Need for robust models for image-based diagnosis which can quickly adapt to novel diseases with few annotated images.

# Related Work

- Several meta-learning techniques have been proposed in the literature and applied for classifying real world scene image datasets.
  - Nichol et al's work 'On first-order meta-learning algorithms.' [1]
  - Snell et al's work 'Prototypical networks for few-shot learning.' [2]
  - Vuorio et al's work 'Multimodal model-agnostic meta-learning via task-aware modulation.' [9]
- There have been a couple of works on meta-learning for skin lesion images.
  - Li et al [7] proposed a difficulty-aware meta-learning method that dynamically monitors the importance of learning tasks and evaluates on ISIC 2018 dataset.
  - Prabhu et al [8] proposed learning a mixture of prototypes for each disease initialized via clustering and refined via an online update scheme.
- G-convolutions [6] greatly improve performance in skin lesion image classification as orientation is not an important feature in such images.

# Contributions

- Propose to use meta-learning for rare disease identification in skin lesion image datasets having long-tailed class distributions and few annotated data samples.
- Explore the gradient based Reptile and metric-learning based Prototypical networks for identifying diseases from skin lesion images.
- Use of Group Equivariant Convolutions (G-Convolutions) improve the network's performance.
- Meta-DermDiagnosis is evaluated on 3 publicly available skin lesion datasets such as ISIC 2018, Derm7pt and SD-198 and compare the classification performance with pre-training as a baseline.
- The proposed meta-learning based disease identification system can also be applied on other medical imaging datasets

# Approach

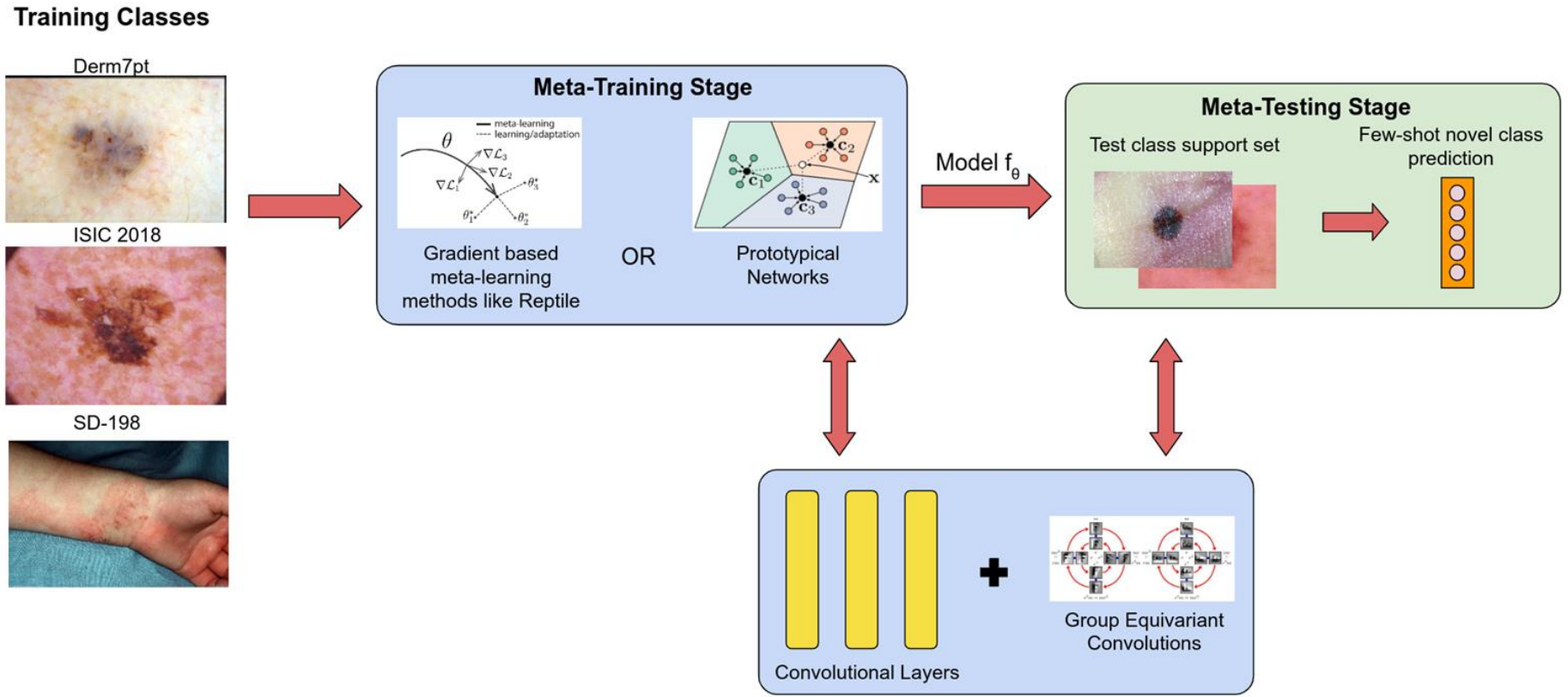



**Figure 3.** Figure showing an overview of the proposed approach ***Meta-DermDiagnosis*** for identification of diseases in skin lesion datasets based on meta-learning techniques Reptile and Prototypical networks.

# Approach

- We select the training classes comprising of common diseases that contain abundant data. Testing classes consist of unseen / rare disease classes with very few examples.
- We use gradient-based Reptile and metric-learning based Prototypical networks along with G-Convolutions (incorporated in the neural networks) for improving few-shot disease classification from skin lesion images.

## Reptile

**Algorithm 1** Reptile [25]

1: Initialize $\boldsymbol{\theta}$, the vector of initial parameters
2: **for** $iteration = 1, 2, \ldots$ **do**
3: Sample task $T$, corresponding to loss $L_T$ on weight vectors $\tilde{\boldsymbol{\theta}}$
4: Compute $\tilde{\boldsymbol{\theta}} = U_T^k(\boldsymbol{\theta})$, denoting $k$ SGD or Adam steps
5: Update $\boldsymbol{\theta} \leftarrow \boldsymbol{\theta} + \epsilon(\tilde{\boldsymbol{\theta}} - \boldsymbol{\theta})$, where $\epsilon$ is the stepsize parameter
6: **end for**

In Reptile algorithm, $U_T^k(\boldsymbol{\theta})$ is the operator (e.g. corresponding to Adam optimizer or SGD) that updates $\boldsymbol{\theta}$ using $k$ mini-batches on data sampled from $T$.

# Reptile

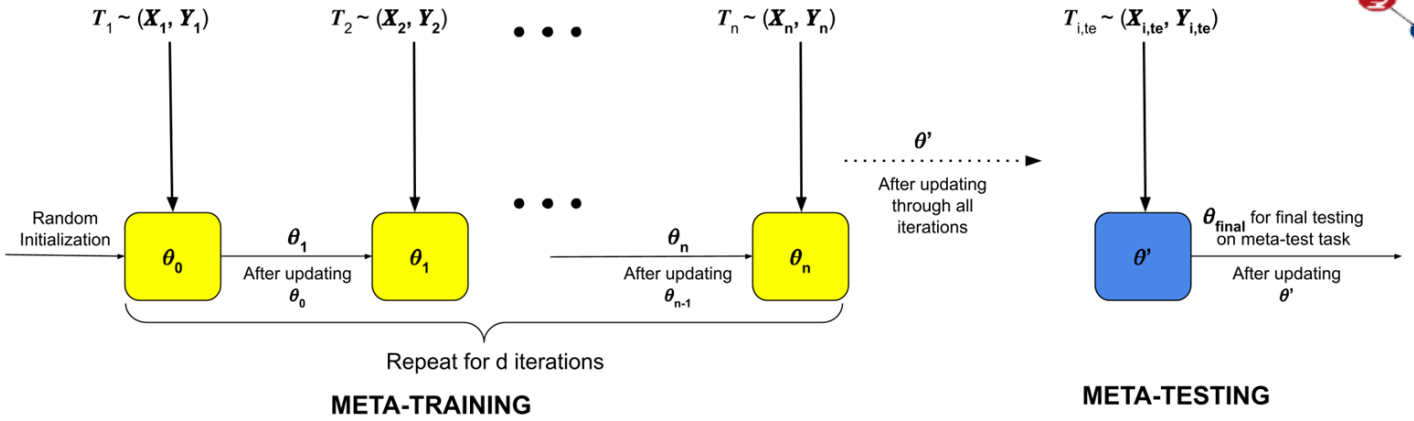



**Figure 4.** Pipeline for gradient-based meta-learning on skin lesion classification.

- 2-way classification tasks for the 3 datasets. For SD-198, 20 train classes and 70 test classes were used, so we also experimented with 4-way classification tasks.
- We query 15 images from the meta-train dataset for each of the classes in a task during the meta-training stage.
- During meta-testing, *k* shot fine-tuning is performed on the meta-trained model. *k* is 1, 3, and 5 indicating 1-shot, 3-shot, and 5-shot respectively.
- The final inference is performed on the entire testing split of the classes in the meta-test task.

# Prototypical Networks

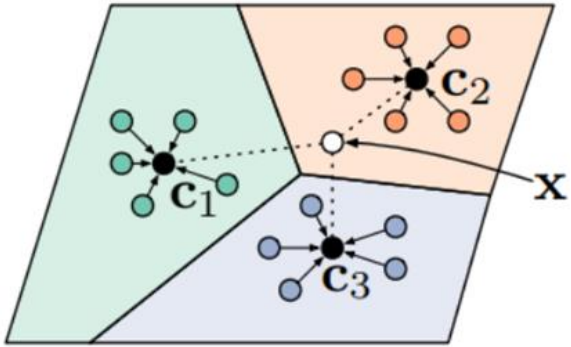



**Figure 5.** Prototypical networks in the few-shot classification. Few-shot prototypes $c_k$ are computed as the mean of embedded support examples for each class.

- Use an embedding function $f_\phi$ to encode each input into a M-dimensional feature vector.
- Let $S_k$ denotes the set of examples labeled with class $k \in C$. A prototype feature vector is defined for each class $k$ as follows:

$$\mathbf{c}_k = \frac{1}{|S_k|} \sum_{(\mathbf{x}_i, y_i) \in S_k} f_\phi(\mathbf{x}_i)$$

- Given a distance function $d$, prototypical networks produce a distribution over classes for a query point $x$ based on a softmax over distances to the prototypes in the embedding space as follows:

$$p_\phi(y = k \mid \mathbf{x}) = \frac{\exp(-d(f_\phi(\mathbf{x}), \mathbf{c}_k))}{\sum_{k'} \exp(-d(f_\phi(\mathbf{x}), \mathbf{c}_{k'}))}$$

## Prototypical Networks

- Trained Euclidean distance-based prototypical networks with the training dataset containing 4, 13, and 20 classes for the ISIC, Derm7pt, and the SD-198 datasets respectively.
- The train-shot is 15, ie. 15 images per class are randomly sampled per episode from the *n* train classes during meta-training, and subsequently the model is trained on these images.
- During meta-testing, 2-way and 4-way classification tasks are created, 1-shot, 3-shot, and 5-shot fine-tuning is performed, and average accuracy and AUC values are computed on the test set.

## Pre-trained Networks (Baseline)

- Involves training a neural network on entire training dataset of all the train classes
- For fine-tuning, 2-way and 4-way classification tasks are created, and 1-shot, 3-shot, 5-shot fine-tuning is performed.
- Average accuracy and AUC is computed on the test dataset of the meta-test tasks created in the previous step.

# Results

Table 1. Performance comparison of AUC (in %) and Accuracy (in %) on the **ISIC 2018** skin lesion dataset for 2-way classification tasks.

| | | Pre-trained | | Reptile | | Prototypical Networks | |
|---|---|---|---|---|---|---|---|
| | | Avg. AUC | Avg. Accuracy | Avg. AUC | Avg. Accuracy | Avg. AUC | Avg. Accuracy |
| **w/o G-Conv** | **1-shot** | 59.7 | 54.8 | 60.3 | 58.0 | **61.6** | **59.3** |
| | **3-shot** | 67.8 | 65.2 | **73.1** | **73.4** | 70.2 | 67.9 |
| | **5-shot** | 72.0 | 71.5 | **79.6** | **76.2** | 75.4 | 73.0 |
| **w/ G-Conv** | **1-shot** | 61.3 | 62.6 | **68.1** | 64.3 | 65.7 | **64.5** |
| | **3-shot** | 72.8 | 69.3 | **81.2** | **76.7** | 75.8 | 73.5 |
| | **5-shot** | 79.1 | 79.4 | **86.8** | **82.1** | 82.9 | 79.7 |

Table 2. Performance comparison of AUC (in %) and Accuracy (in %) on the **Derm7pt** skin lesion dataset for 2-way classification tasks.

| | | Pre-trained | | Reptile | | Prototypical Networks | |
|---|---|---|---|---|---|---|---|
| | | Avg. AUC | Avg. Accuracy | Avg. AUC | Avg. Accuracy | Avg. AUC | Avg. Accuracy |
| **w/o G-Conv** | **1-shot** | 56.9 | 58.4 | 59.7 | 60.2 | **60.6** | **62.5** |
| | **3-shot** | 62.1 | 60.7 | 64.1 | **65.7** | **65.8** | 63.9 |
| | **5-shot** | 66.6 | 64.9 | **71.4** | **70.5** | 68.2 | 66.7 |
| **w/ G-Conv** | **1-shot** | 60.8 | 59.5 | 62.1 | 61.8 | **63.7** | **64.1** |
| | **3-shot** | 62.6 | 62.3 | **68.7** | **69.9** | 65.3 | 66.8 |
| | **5-shot** | 69.8 | 65.2 | **77.2** | **76.9** | 72.8 | 69.5 |

Table 3. Performance comparison of AUC (in %) and Accuracy (in %) on the **SD-198** skin lesion dataset for 2-way classification tasks.

| | | Pre-trained | | Reptile | | Prototypical Networks | |
|---|---|---|---|---|---|---|---|
| | | Avg. AUC | Avg. Accuracy | Avg. AUC | Avg. Accuracy | Avg. AUC | Avg. Accuracy |
| **w/o G-Conv** | **1-shot** | 56.4 | 55.7 | **64.1** | **63.0** | 59.4 | 59.8 |
| | **3-shot** | 65.3 | 60.7 | **77.4** | **72.9** | 70.6 | 66.6 |
| | **5-shot** | 77.9 | 73.6 | **84.6** | **80.4** | 80.7 | 78.3 |
| **w/ G-Conv 2-way** | **1-shot** | 57.4 | 56.9 | **68.6** | **65.3** | 62.9 | 64.5 |
| | **3-shot** | 70.2 | 69.1 | **79.1** | **75.8** | 74.5 | 72.1 |
| | **5-shot** | 84.2 | 76.5 | **89.5** | **83.7** | 85.6 | 80.2 |

- In some 1-shot learning cases like for ISIC and Derm7pt datasets, the prototypical networks perform slightly better than Reptile.
- For slightly higher number of samples, Reptile outdoes prototypical networks.
- Performance of meta-learning and baseline pre-training: 5-shot > 3-shot > 1-shot.
- Use of G-convolutions improves the network's performance on all 3 datasets as they make the neural network equivariant to image transformations.

# References


[1] Nichol, Alex, Joshua Achiam, and John Schulman. "On first-order meta-learning algorithms." *arXiv preprint arXiv:1803.02999* (2018).

[2] Snell, Jake, Kevin Swersky, and Richard Zemel. "Prototypical networks for few-shot learning." *Advances in neural information processing systems*. 2017.

[3] Codella, Noel, 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).

[4] Kawahara, Jeremy, et al. "Seven-point checklist and skin lesion classification using multitask multimodal neural nets." *IEEE journal of biomedical and health informatics* 23.2 (2018): 538-546.

[5] Sun, Xiaoxiao, et al. "A benchmark for automatic visual classification of clinical skin disease images." *European Conference on Computer Vision*. Springer, Cham, 2016.

[6] Cohen, Taco, and Max Welling. "Group equivariant convolutional networks." *International conference on machine learning*. 2016.

[7] Li, Xiaomeng, et al. "Difficulty-aware Meta-Learning for Rare Disease Diagnosis." *arXiv preprint arXiv:1907.00354* (2019).

[8] Prabhu, Viraj, et al. "Few-Shot Learning for Dermatological Disease Diagnosis." *Machine Learning for Healthcare Conference*. 2019.

[9] Vuorio, Risto, et al. "Multimodal Model-Agnostic Meta-Learning via Task-Aware Modulation." *Advances in Neural Information Processing Systems*. 2019.

# Thank You!

Contact us:

TCS Research, New Delhi
India

kushagra.mahajan@tcs.com