Semi-supervised few-shot learning for medical image segmentation

Abdur R Fayjie, Reza Azad,.., Ismail Ben Ayed, and Jose Dolz

Presented by: Kaveh Moradkhani

Neuro-iX

02

Importance of Semantic Segmentation:

- ❖ Crucial for treatment, diagnosis, and monitoring of diseases
- ❖ Automation in this field has been widely researched
- ❖ Recent advancements in deep learning are enhancing capabilities

Role of Convolutional Neural Networks (CNNs):

- ❖ Achieved top performance in medical image segmentation
- ❖ Applications include brain tissue, heart structures, and abdominal organs
- ❖ High-performing models require large labeled datasets

Challenges in Traditional Methods

- **Need for Large Labeled Datasets**

  High-quality segmentation requires extensive annotated data

- **Limitations**

  Difficult, time-consuming, and costly to obtain, especially in medical contexts

  Risk of overfitting when data is limited

- **Main Challenge**

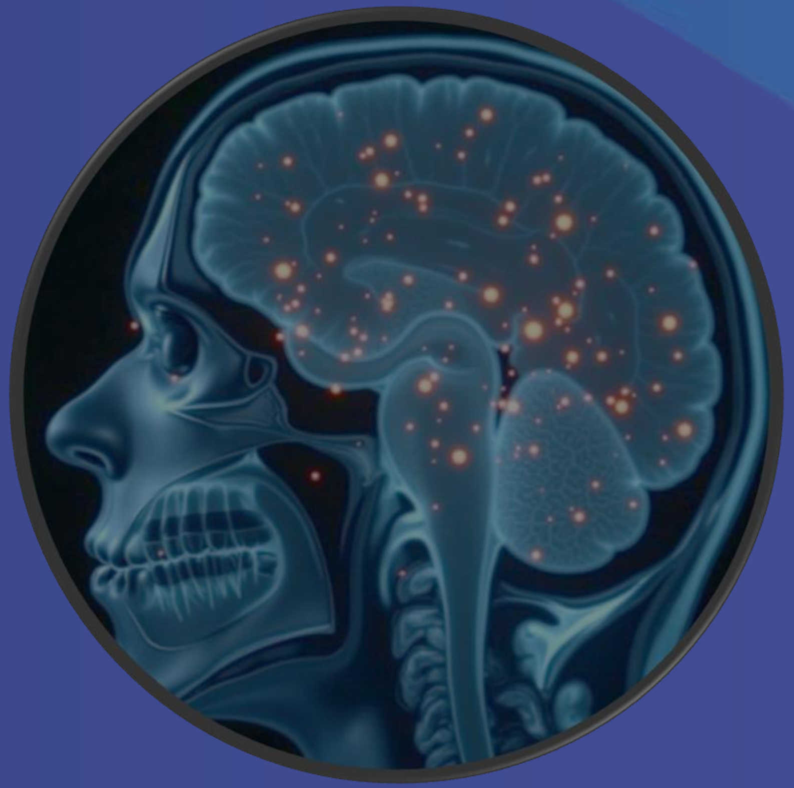
  How to learn effectively from limited labeled samples

04

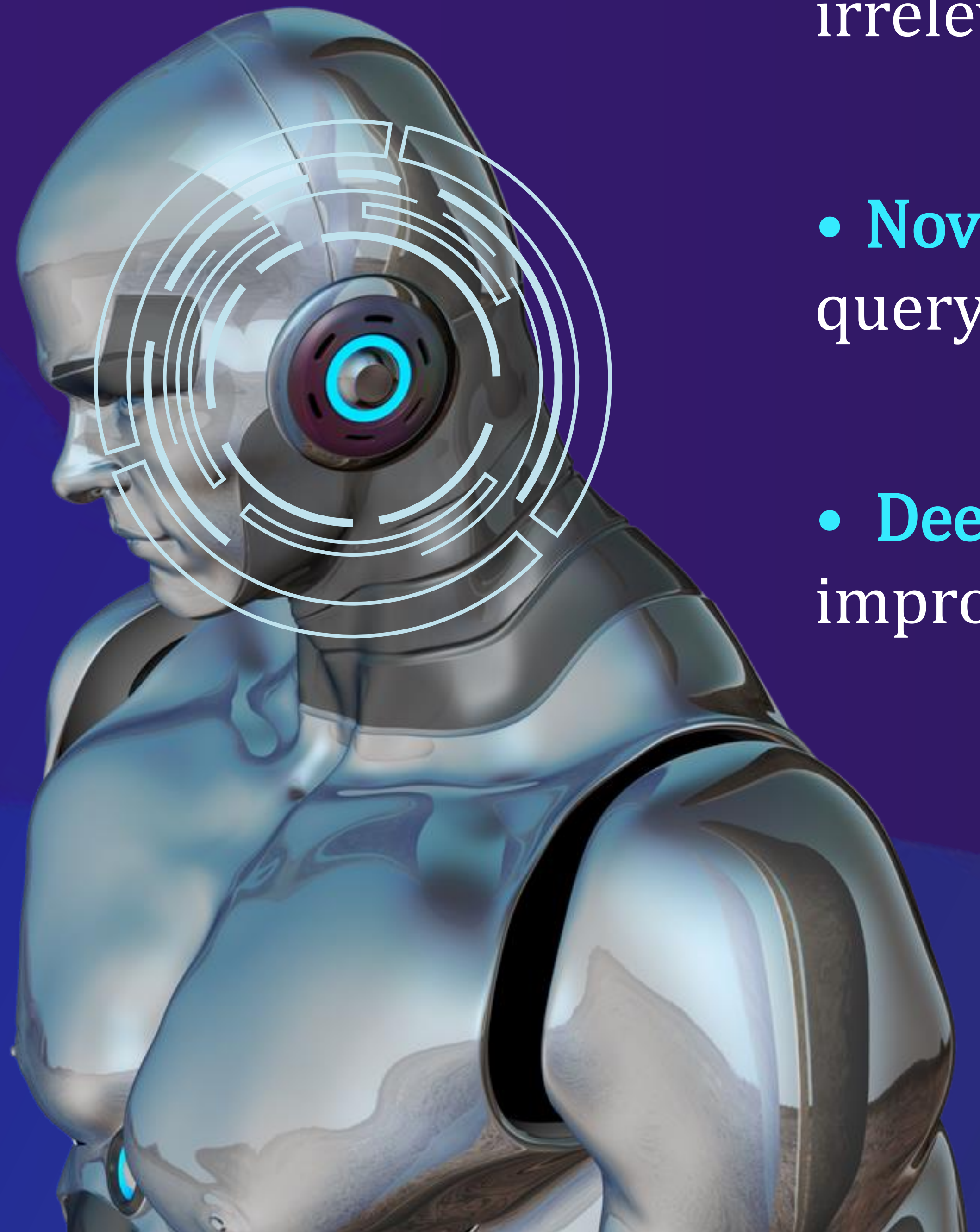
Main Goal

❖ Propose a novel framework for few-shot learning in semantic segmentation

❖ Integrate unlabeled data to improve performance



Techniques of Previous Methods to Improve Generalization



- **Mask Average Pooling Strategy:** Focuses on relevant features by ignoring irrelevant ones based on support masks
- **Novel Prototype Alignment:** Aligns features from support images with those in query images, enhancing class recognition
- **Deep Attention Mechanisms:** Learns to focus on important parts of images, improving label propagation from support to query images

next slide

Limitations of Current Approaches

- **Lack of Unlabeled Data Utilization**:

  Many methods overlook unlabeled data,

  missing valuable information that could enhance performance

Introduction to Few-Shot Learning

- Efficient alternative to traditional supervised learning
- Trains on a few labeled images (support images)
- Uses knowledge from support images for segmentation of new classes (query images)

Advantages	Disadvantages
Reduces dependency on large labeled datasets	Limited Generalisation
Improves adaptability to new tasks	Sensitivity to Support Set Composition
Useful in domains with limited data availability (e.g., medical imaging)	Difficulty with Out-of-Distribution Data
	Overfitting to Support Set

next slide

Proposed Innovations in Few-Shot Learning

- **Investigating Unsupervised Data:**
  - The role of unsupervised data using surrogate tasks to enhance few-shot learning for medical image segmentation
- **Episodic Training Paradigm:**
  - applying the episodic training paradigm to the medical context
- **Improving Model Performance:**
  - Incorporating unlabeled images into surrogate tasks aims to improve generalization capabilities
- **Episodic Training Structure:**
  - k labeled examples: Small number of support images
  - n novel categories: Different new categories for the model to learn
- **Purpose of Episodes:**
  - This approach exposes the model to diverse tasks, reducing over-specialization and enhancing generalization to new examples

Datasets

- **Training Set (D_{train}):**

This is where the model learns from a collection of images and their corresponding labels. It consists of pairs (X_i^t, Y_i^t)

- **Support Set ($D_{Support}$):**

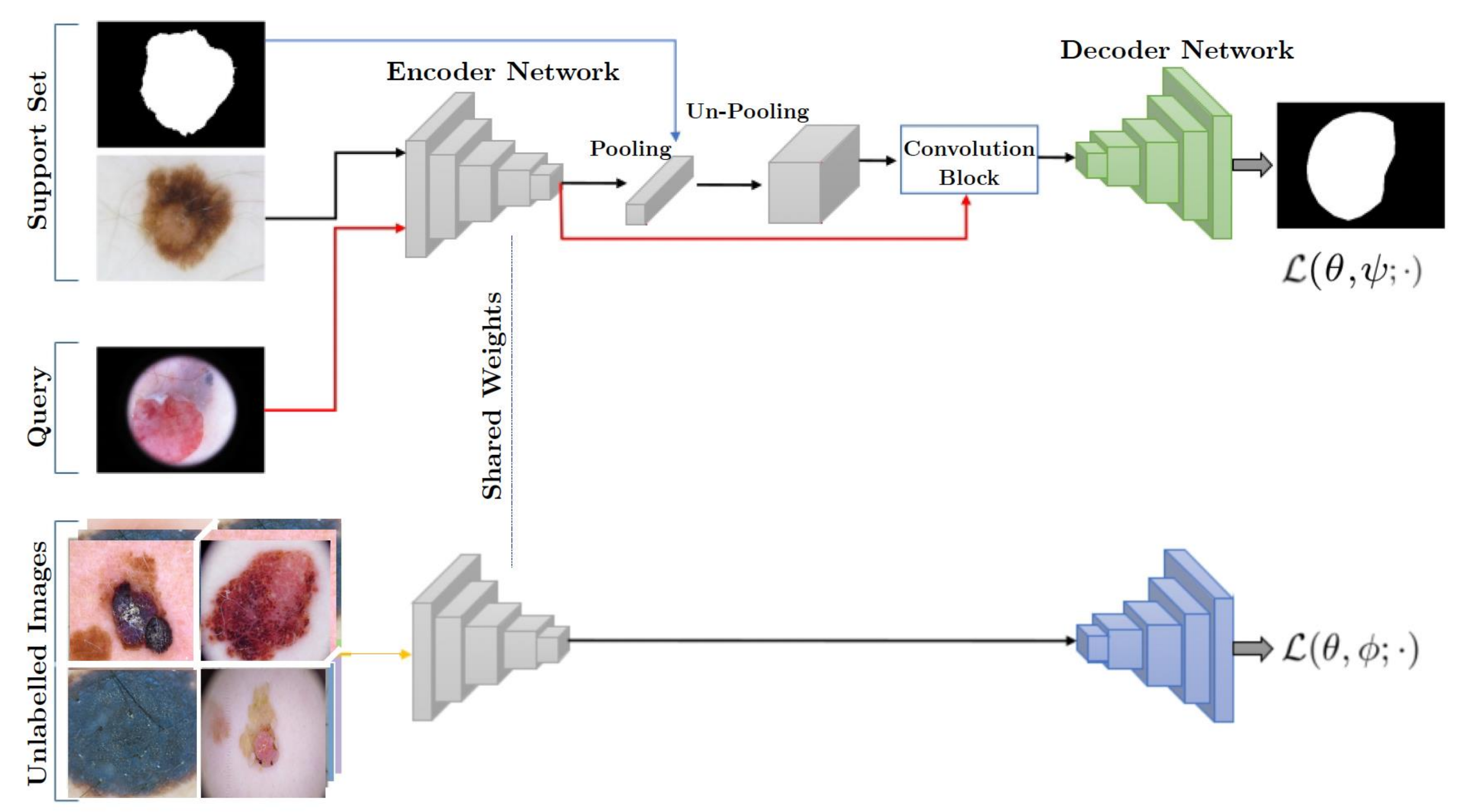
This is a small set of labeled images that the model will use as references for learning new classes. It consists of pairs (X_i^s, Y_i^s)

- **Test Set (D_{test}):**

This set contains images that the model has never seen before, and it will predict the segmentations for these images. It consists of X_i^q

Goal of Few-Shot Learning: The objective is to train a neural network $f_{(\theta,\psi)}(.)$ to segment new classes c that are not present in the training set C_{train}. The model does this using k reference from the support set $D_{support}$

Model Overview



- Using a CNN to convert both the support and query images into a lower-dimensional feature space

  $f_s \in \mathbb{R}^{w' \times H' \times M}$ for the support images

  $f_q \in \mathbb{R}^{w' \times H' \times M}$ for the query images

- Class Prototype Calculation
  - Computes class prototypes using mask average pooling on the feature representation of support images:

$$P_s = \frac{1}{|\widetilde{Y_s^t(c)}|} \sum_{i=1}^{w' \times H'} f_s \widetilde{Y_s^t(c)}$$

  - $\widetilde{Y_s^t(c)}$ is a down-sampled version of the binary mask for the support image, indicating which pixels belong to the class

- Optimization
  - Model parameters (θ, ψ) are optimized by minimizing the difference between predicted segmentation mask $\widetilde{Y_q^t(c)}$ and actual segmentation mask $Y_q^t(c)$ $\quad \widetilde{Y_q^t}(c) = f_{(\theta,\psi)}(D_{train}^S, X_q^t)$

- Evaluation
  - During testing, the model $f_{(\theta,\psi)}(.)$ is evaluated on the test set D_{test} using the k images from the support set $D_{support}$

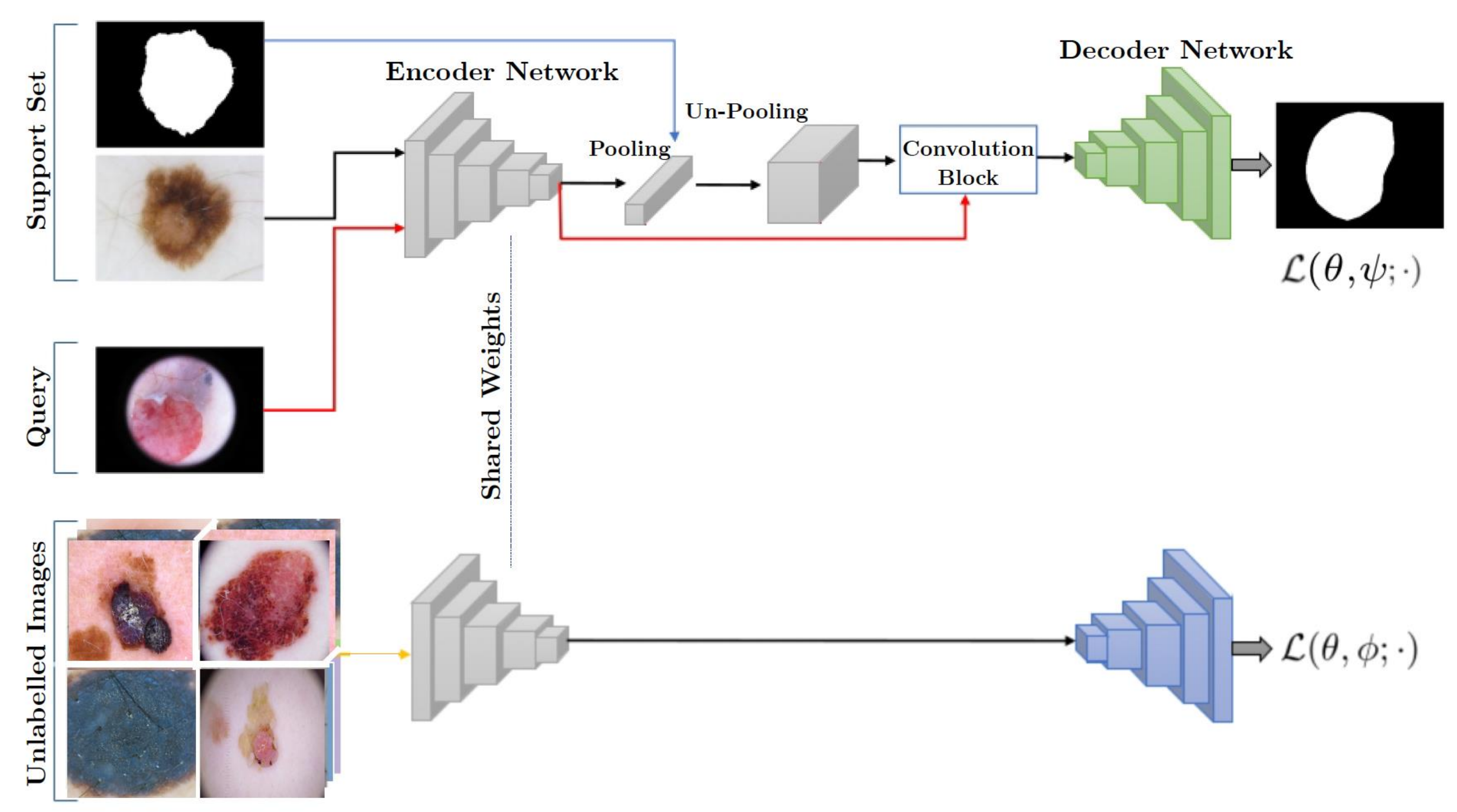
❖ Key Challenge:

Training the model's feature extractor to recognize useful image features for new categories with limited labeled examples

❖ Proposed Solution:

Integrate Auxiliary Loss: Introduce an auxiliary loss during training defined as $\mathcal{L}(\theta, \emptyset; .)$

$\emptyset$: Parameters related to the surrogate task



❖ Task Selection:

Image Denoising: Chosen for its suitability in medical imaging

Objective: Remove noise from images to recover clear versions

❖ Use of Unlabeled Data:

Additional dataset of unlabeled images: $D_u^{train} = \{X_u^i\}_{i=1}^{N_u}$

Random noise is added to create corrupted images Y_i^u

❖ Training Objective:

Train the model to map noisy images back to their original forms using the **cross-entropy loss function**:

$$\mathcal{L}_{sur}(\theta, \psi; Y^u) = -\frac{1}{N_u} \sum_{i=1}^{N_u} \sum_{j=1}^{H \times W} X_{i,j}^u \log \widetilde{Y_{i,j}^u}$$

$\widetilde{Y_i^u} = f_{(\theta,\psi)}(Y_i^u)$ represents the denoised version of the noisy image

The overall goal training is to optimize two main components

Few-Shot Segmentation Loss ($\mathcal{L}_{few}(\theta, \psi; D_{train})$):

This is related to the parameters θ and ψ used in the few-shot segmentation task

Surrogate Task Loss ($\mathcal{L}_{sur}(\theta. \phi; Y^u)$):

This focuses on the parameters θ related to the encoder and the parameters ϕ for the surrogate task

❖ The training process aims to minimize the combined objective function:

$$min_{\theta,\psi,\phi}\mathcal{L}_{few}(\theta, \psi; D_{train}) + \lambda\mathcal{L}_{sur}(\theta, \phi; Y^u)$$

λ is a weighting factor that indicates how much importance should be given to the surrogate task compared to the few-shot segmentation task

Datasets Used

FSS-1000 Dataset

A large-scale few-shot segmentation dataset with 1000 classes, each containing 10 images and pixel-level annotations-240 classes are used for surrogate tasks

ISIC Dataset

2594 dermoscopic images with masks for skin lesions. 1594 images are used for surrogate tasks

PH^2 Dataset

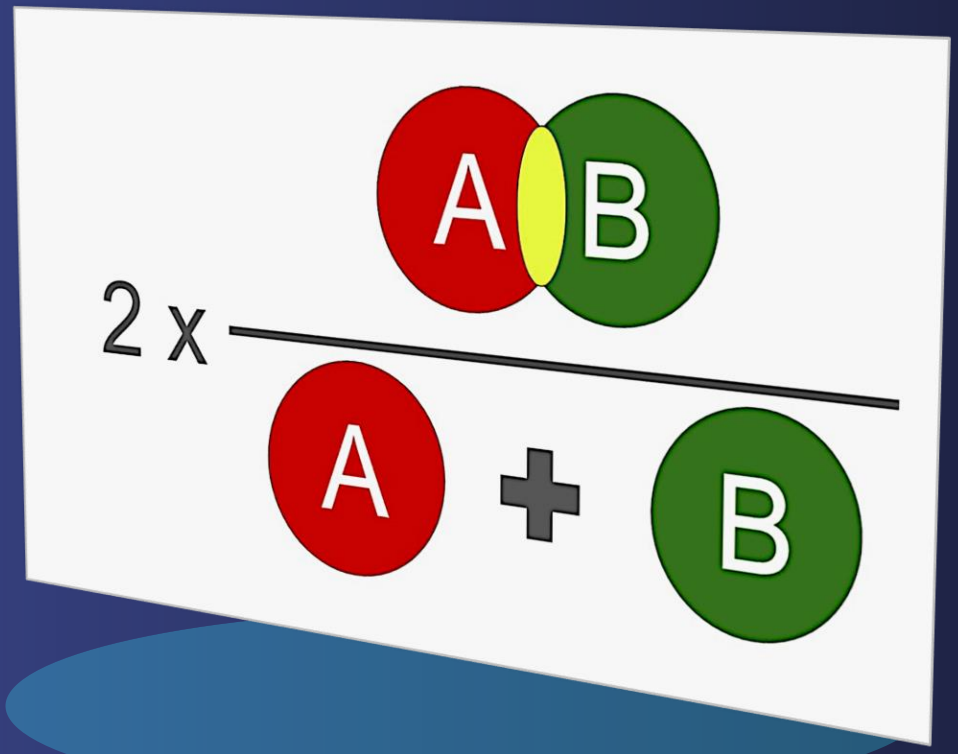
200 images of melanocytic lesions used exclusively for testing

Performance Metric:

$$\mathrm{DSC} = \frac{2|A \cap B|}{|A| + |B|}$$

A and B represent the two segmentation masks being compared

The numerator calculates the intersection (overlap) between the two masks, while the denominator sums the total number of pixels in both masks



Quantitative Results

❖ **1-Shot Scenario:**

Effectiveness of few-shot learning, Compared to standard batch-wise training (Regular), episodic training shows an improvement of nearly 6% on ISIC and PH^2datasets. model trained on FSS-1000 and directly tested on ISIC and PH^2

Leveraging unsupervised data via surrogate tasks improves performance by 6-7% over the few-shot baseline

Compared to the upper bound (same dataset training/testing), the proposed model show approximately 15% difference in results on the PH^2 dataset, but only segments 1 target image

Additional Samples:

Increasing the number of samples in episodes leads to better results, indicating effective extraction of semantic information from unlabeled data

❖ **5-Shot Scenario:**

When integrating 10 additional images, the model improves by 3-4% compared to the baseline without surrogate tasks

Improvement is more significant with fewer labeled samples (1-shot vs. 5-shot)

Model	Additional Samples	ISIC	PH^2
Regular (*Lower-bound*)	—	48.32	64.72
Few-shot	—	54.07	68.13
Few-shot + unlabeled (*Denoising*) ***(Ours)***	5	61.38	74.12
	10	**61.40**	74.67
	20	60.79	**74.77**
Regular (*Upper-bound*)	—	86.65	89.94

Quantitative Results of 1-Shot Settings on ISIC and PH^2 Datasets

Model	ISIC	PH^2
Few-shot	59.63	71.15
Few-shot + Unlabeled (*Denoising*) ***(Ours)***	**62.40**	**75.54**

Quantitative Results of 5-Shot Settings on ISIC and PH^2 Datasets

Qualitative Results

Batch-Wise Training (Right Column):

Segmentation results are unsatisfactory

Notably over-segmenting the target areas

Episodic Training (Third Column):

Shows improved capability to learn general features

Reduced false positives (isolated pixels and over-segmentations)

Few-Shot Learning with Unlabeled Data (Fourth Column):

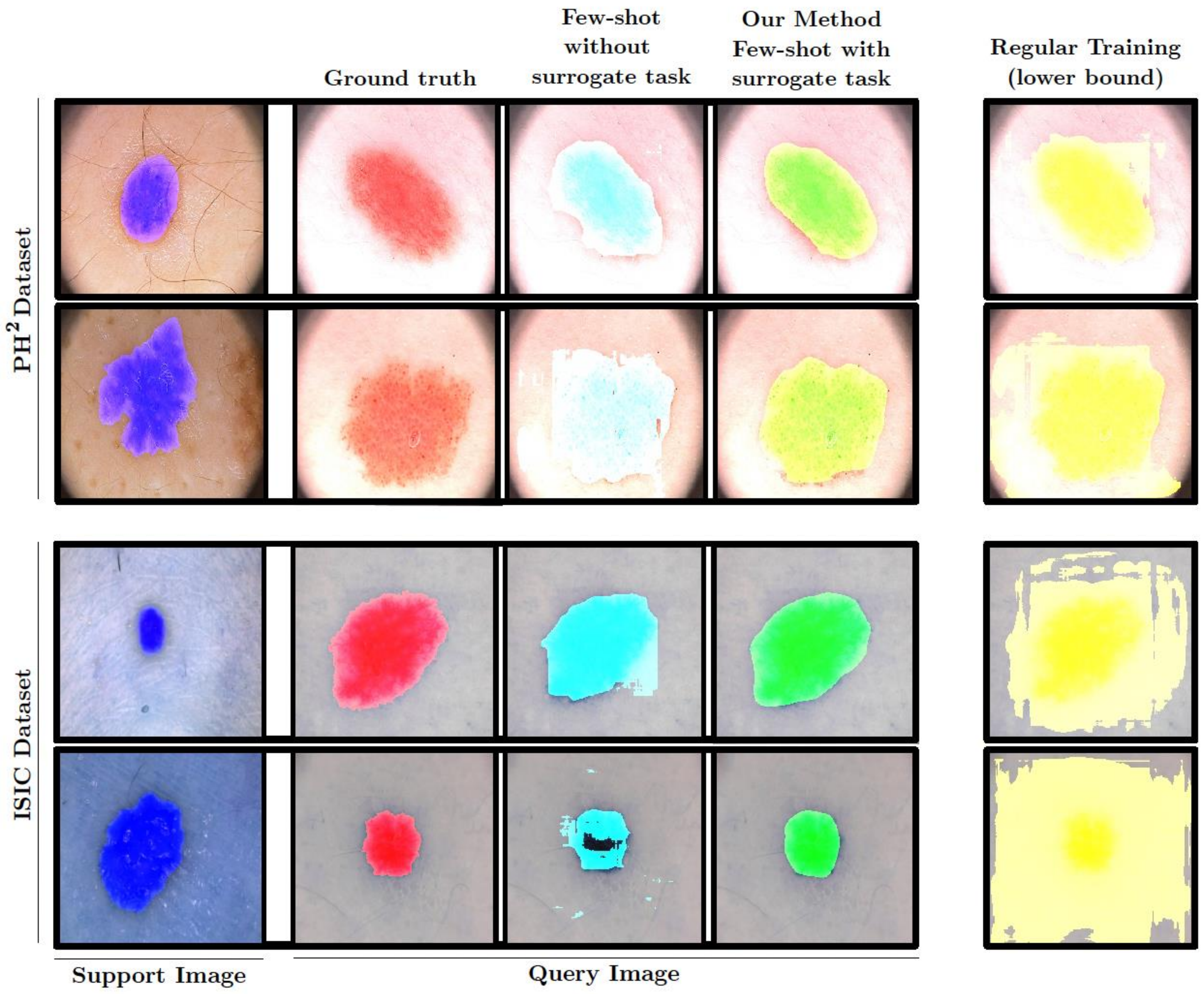
Further enhances representation power

Produces richer and more generic features

Impact on Segmentation:

Improved segmentation with more true positives (correctly identified target areas)

Better segmentation results compared to the batch-wise approach



Conclusion

❖Integration of Episodic Training:

First attempt to apply episodic training in few-shot semantic segmentation for medical images

❖Surrogate Tasks:

Investigated the use of surrogate tasks, inspired by self-supervised learning, to enhance few-shot segmentation

Auxiliary tasks like image denoising help the encoder learn rich, transferable features

❖Improved Performance:

Enriched features lead to better representations and generalization to unseen tasks

Experiments on public skin cancer datasets demonstrate significant performance improvements by leveraging unlabeled images

Thank You!

Kaveh.Moradkhani.1@ens.etsmtl.ca

https://kaveh-moradkhani.github.io/