

Comparing Pretraining Strategies for Intra-Oral Image Analysis Under Limited Annotation Budgets

Nicolas MARTIN ^{1,2}, Philippe MULHEM ², Jean-Pierre CHEVALLET ²

¹ PEEKTORIA, ² Univ. Grenoble Alpes, CNRS, Grenoble INP, LIG

Motivation

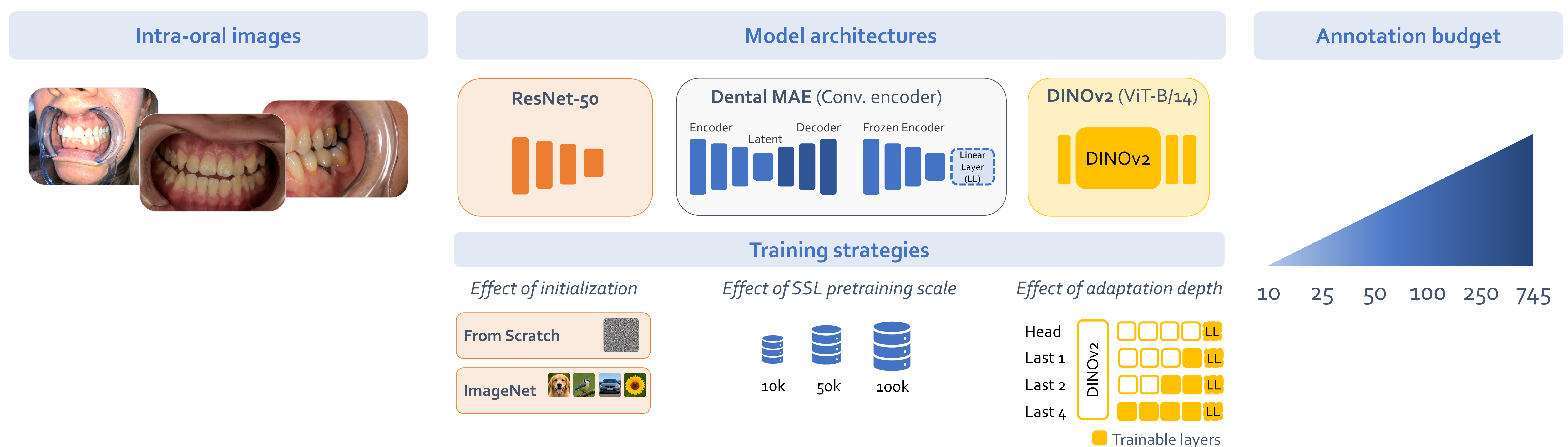
Expert annotations are scarce and costly for intra-oral image analysis.

Transfer learning, foundation models, and domain-specific Self-Supervised Learning (SSL) offer alternative strategies for data-efficient learning.

Their effectiveness across annotation budgets remains unclear.

Experimental setup

We evaluate how initialization, domain-specific SSL pretraining scale, and DINOv2 adaptation depth interact with the amount of labeled data available.



Results

Best performance: DINOv2 - last 4 blocks (Macro-F1 = 0.683 ± 0.034)

- Frozen DINOv2 competitive at low-label budgets
- Deeper adaptation benefits from more labels
- Scaling Dental MAE pretraining (10k $\rightarrow$ 100k) showed no consistent gain

Take-home message

DINOv2 provides strong representations for intra-oral images, with partial adaptation achieving the best performance as more labels become available.

Future works:

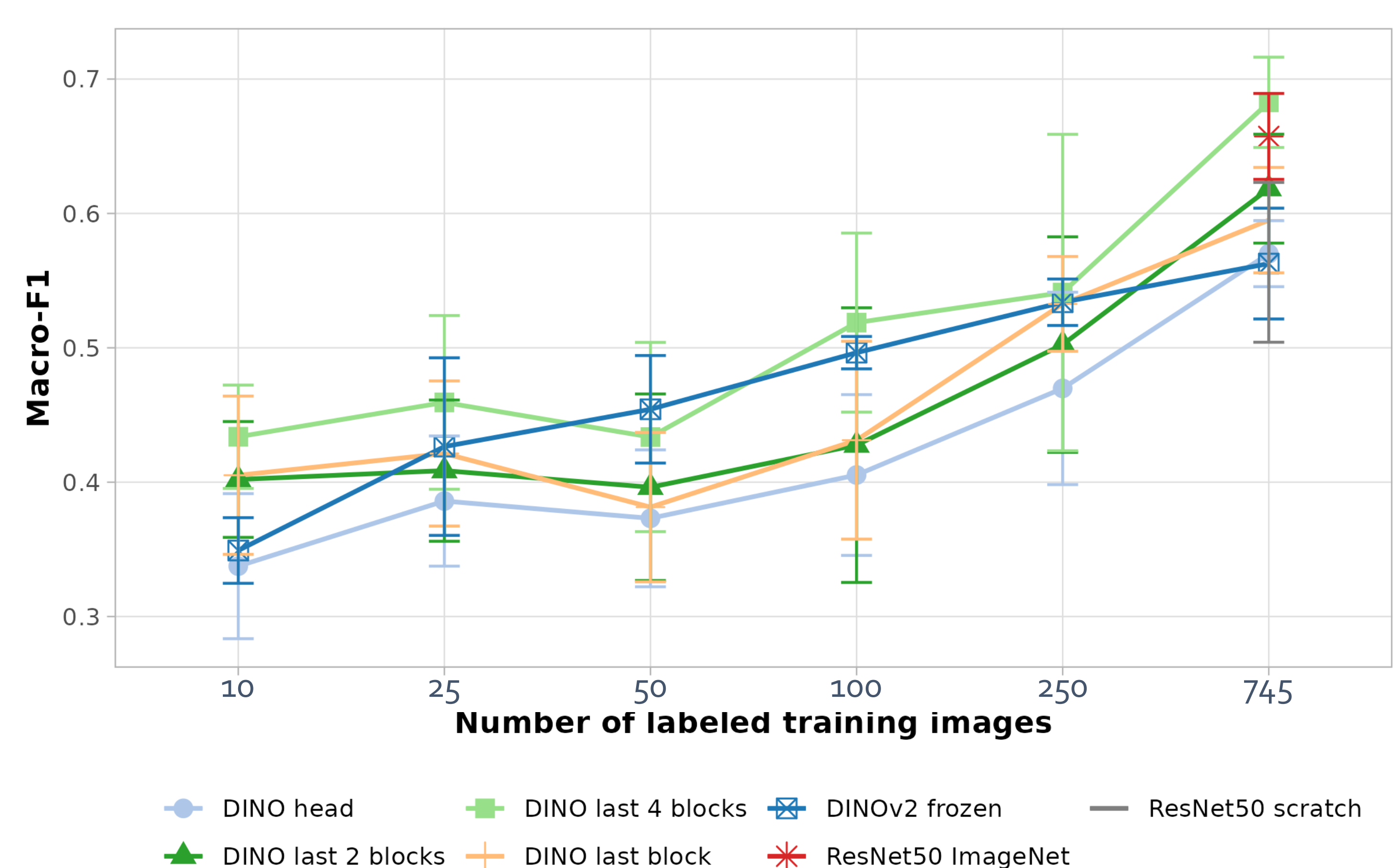
- Scale domain-specific SSL to larger, capacity-matched architectures
- Continue SSL pretraining of foundation models on intra-oral images

Research Question

Which representation learning strategy is most effective for intra-oral image analysis under limited annotation budgets?

Intra-oral Datasets (142k images)

1. *Domain-specific SSL pretraining*
 - 140k unlabeled intra-oral photographs
2. *Periodontitis classification*
 - 931 patients / images
3. *Public Gingivitis dataset (external validation)*
 - 1,096 images



External validation

Method	Macro-F1
ResNet50 ImageNet	.377
DINOv2 frozen	.337
Dental MAE-50k	.160

Details:

