#5658

TOWARDS AN OPTIMAL (SELF)-SUPERVISED LEARNING PARADIGM FOR DIABETIC RETINOPATHY CLASSIFICATION

Christopher Clark, MS[1], Scott Kinder, MS[1], Advaith Veturi, MS [1], Steve McNamara, OD[1] , Benjamin Bearce, MS[1], Ramya Gnanaraj, MD[1], Niranjan Manoharan, MD[1], Talisa de Carlo, MD[1], Naresh Mandava, MD [1], Malik Kahook, MD[1], Praveer Singh, PhD[1], Jayashree Kalpathy-Cramer, PhD[1]

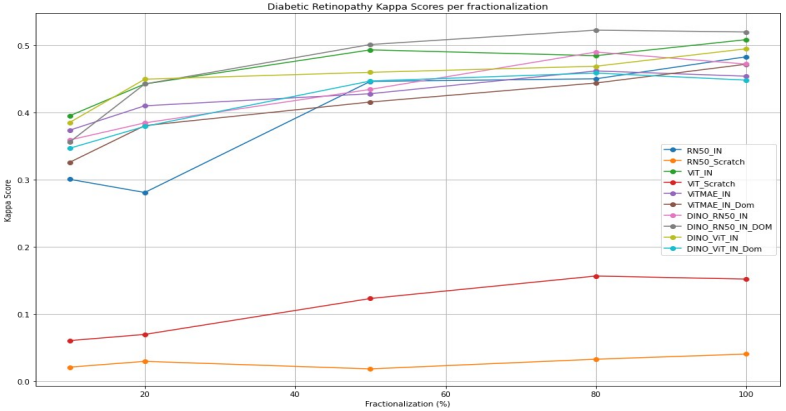
1Ophthalmology, University of Colorado Anschutz School of Medicine, Aurora, CO

Purpose

- Currently, no evaluation study is exploring the most optimal strategy for training an Artificial Intelligence (AI) based diabetic retinopathy classification model.
- To remedy this, in this study, we compare existing **Supervised Learning (SL)** approaches with various **Self-Supervised Learning (SSL)** techniques, for different architecture backbones, pretraining strategies, and generative vs contrastive SSL methodologies.

Methods

- SSL with **generative (MAE)** vs **contrastive (DINO)** pretext learning paradigms.
- **Pretext fine-tuning** with just ImageNet vs adding domain-related data.
- **DINO CNN** vs **transformer** backbone architecture.



Methods Cont.

- **SL fine-tuning** with ImageNet vs from scratch on domain-specific data.
- The role the **amount of data** plays in each algorithm's performance (10, 20, 50, 80, and 100% of the downstream, classification data)

Results

- The **best model** was a **DINO CNN (ResNet50)** model pretext fine-tuned on domain-specific data.
- In the **DINO paradigm without domain-specific** pretext fine-tuning, the transformer (ViT) performed better than the CNN except for 80% fractionation. With **domain-specific pretext fine-tuning**, the opposite occurred.
- In the **SSL MAE paradigm**, pretext fine-tuning with domain-specific data only surpassed the non-domain-specific fine-tuned at 100% fractionation.
- In a **strictly SL** paradigm, fine-tuning on ImageNet greatly increased performance.
- Overall**, self-supervised learning** can improve diabetic retinopathy deep learning classification algorithms, though care must be taken to choose the right pretext training paradigm.