Self-supervised pre-training with diffusion model for few-shot landmark detection in x-ray images
–Supplementary Materials–

1. Performance evaluation on the Chest dataset

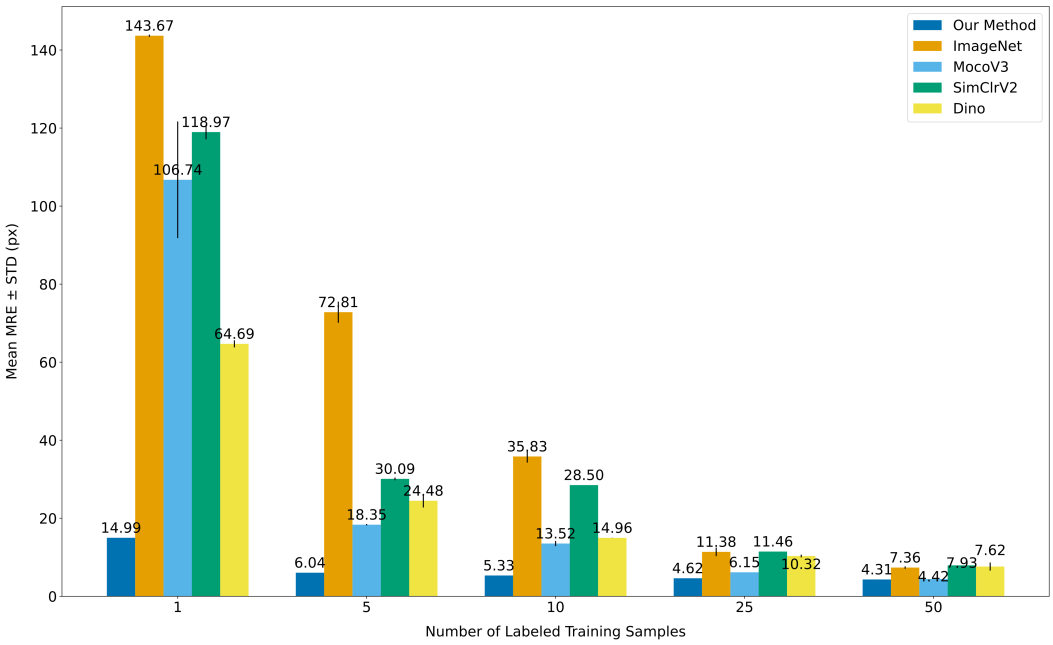

Figure 1. Comparison of landmark detection performance (Mean MRE in pixels) on the Chest dataset using various pre-training methods with limited labeled data ($1 - 50$ samples). Our DDPM self-supervised approach consistently outperforms ImageNet supervised pre-training and state-of-the-art self-supervised methods (MoCoV3, SimCLRV2, DINO) across all sample sizes. The performance gains are most pronounced in low-data regimes: with one labeled sample, our method achieves an MRE of 14.99 px, representing a 76.8% improvement over the best alternative method (DINO at 64.69 px) and an 89.6% improvement over ImageNet (143.67 px). While the performance gap decreases as more labeled samples are added, our method still maintains the top spot, achieving the lowest MRE even with 50 samples.

2. Performance evaluation on the Cephalometric dataset

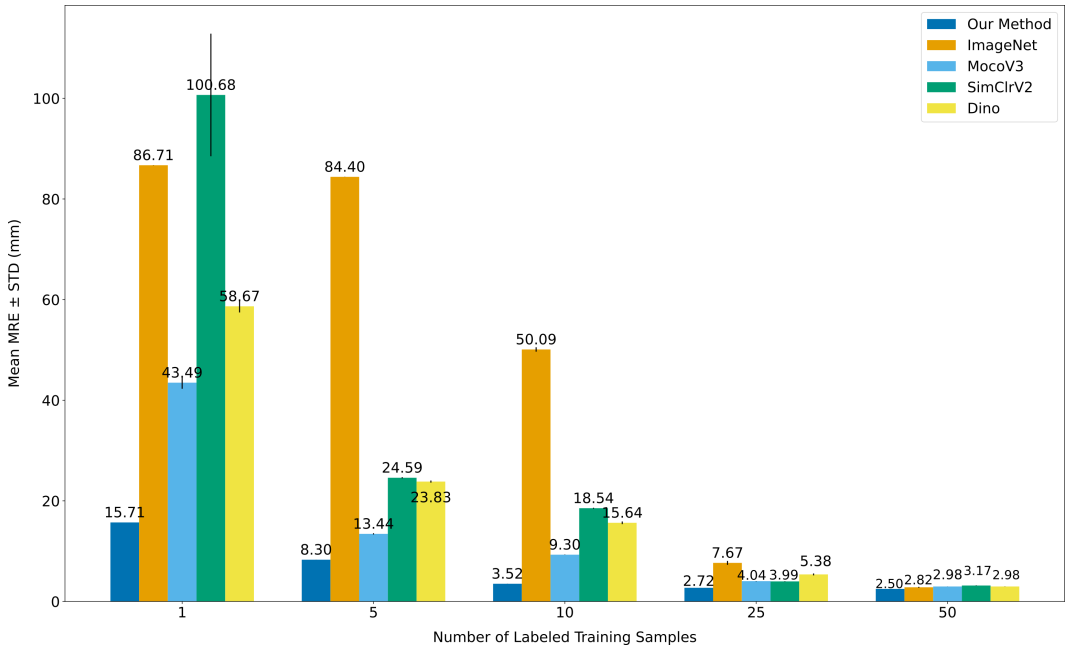

Figure 2. Comparison of landmark detection performance (Mean MRE in mm) on the Cephalometric dataset using various pre-training methods with limited labeled data ($1 - 50$ samples). Our DDPM self-supervised approach consistently outperforms ImageNet supervised pre-training and state-of-the-art self-supervised methods (MoCoV3, SimCLRV2, DINO) across all sample sizes. The performance gains are most pronounced in low-data regimes: with one labeled sample, our method achieves an MRE of 15.71 mm, representing a 63.88% improvement over the best alternative method (MoCoV3 at 43.49 mm) and an 81.88% improvement over ImageNet (86.71 mm). While the performance gap decreases as more labeled samples are added, our method still maintains the top spot, achieving the lowest MRE even with 50 samples.

3. Performance evaluation on the Hand dataset

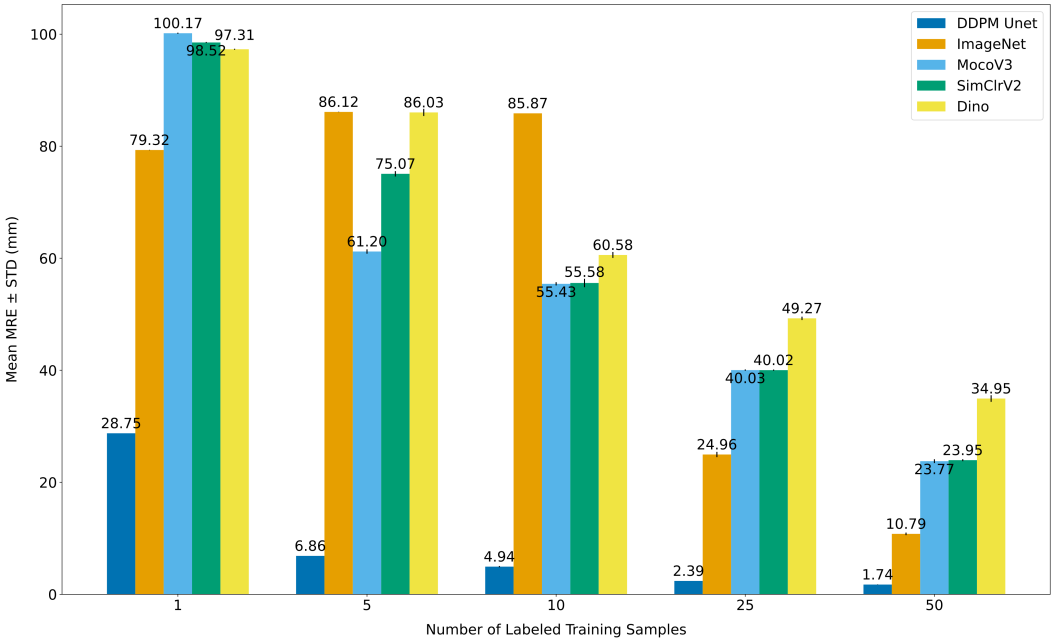

Figure 3. Comparison of landmark detection performance (Mean MRE in mm) on the Hand dataset using various pre-training methods with limited labeled data ($1 - 50$ samples). Our DDPM Unet approach consistently outperforms ImageNet supervised pre-training and state-of-the-art self-supervised methods (MoCoV3, SimCLRV2, DINO) across all sample sizes. The performance gains are most pronounced in low-data regimes: with one labeled sample, our method achieves an MRE of 28.75 mm, representing a 63.8% improvement over ImageNet (79.32 mm) and a 70.5% improvement over the next best self-supervised method (DINO at 97.31 mm). While the performance gap decreases as more labeled samples are added, our method still maintains the top spot, achieving the lowest MRE even with 50 samples.