

Pre-training Vision Transformers with Very Limited Synthesized Images

Ryo Nakamura^{1,2}, Hirokatsu Kataoka¹, Sora Takashima^{1,3}, Edgar Josafat Martinez Noriega^{1,3}, Rio Yokota^{1,3}, Nakamasa Inoue^{1,3} ¹ National Institute of Advanced Industrial Science and Technology (AIST), ² Fukuoka University, ³ Tokyo Institute of Technology

ViT pre-training can be done with the listed 1,000 images!