



Video Pretraining Advances 3D Deep Learning on Chest CT Tasks

Alexander Ke, Shih-Cheng Huang, Chloe P O'Connell, Michal Klimont, Serena Yeung, Pranav Rajpurkar

alexke@cs.stanford.edu, mschuang@stanford.edu, chloe.p.oconnell@gmail.com, mklimont@stanford.edu, syyeung@stanford.edu, pranav_raipurkar@hms.harvard.edu



Contributions

- We study the impact of (1) video pretraining, in-domain pretraining, and sequential pretraining for a (2) broad, representative universe of models (seven 3D and three 2D) on (3) two large-scale public datasets of chest CTs for PE detection and lung nodule detection, across (4) three dataset sizes.
- Our direct and original comparisons of video with in-domain and sequential pretraining help disentangle the effect of video pretraining from other pretraining procedures.
- Our experiments illuminate how pretraining's benefits scale with dataset size and how pre- and post-training methods interact with performance, which are especially important in the small data regimes of medicine.

2D vs. 3D Comparison

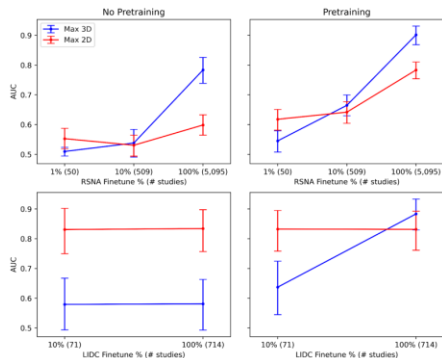


Figure 1: With video pretraining, 3D models outperform 2D models on less data

Method

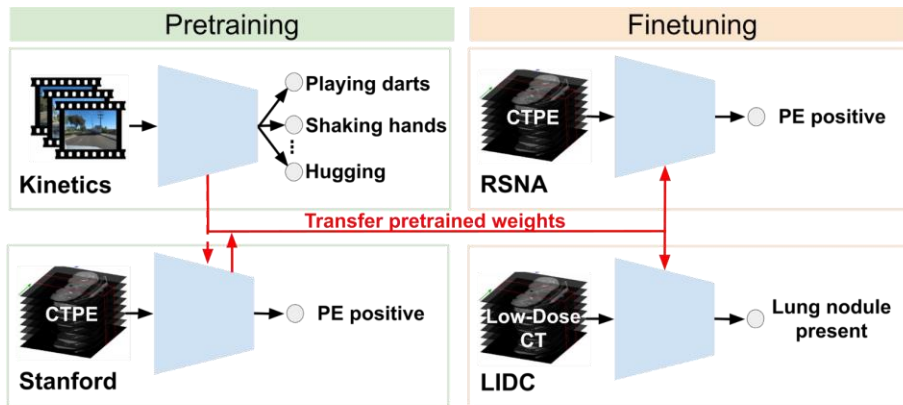
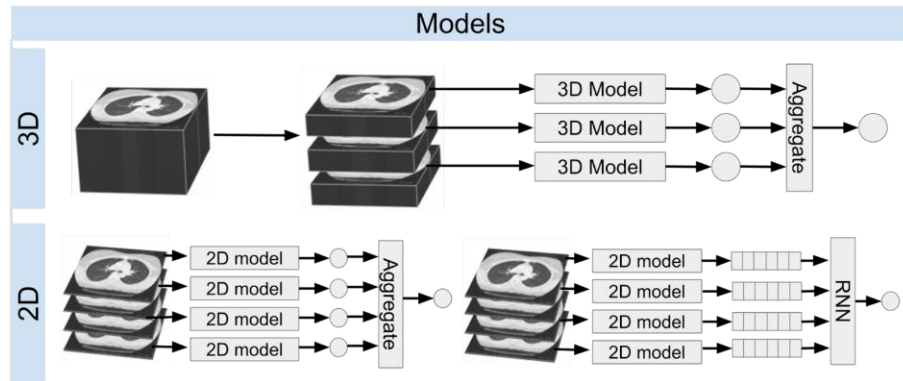


Figure 2: Visual summary of our methods.



Pretraining Data Comparisons

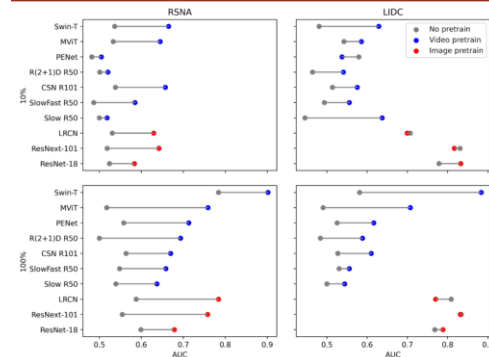


Figure 3: Video pretraining consistently improves 3D models' AUC

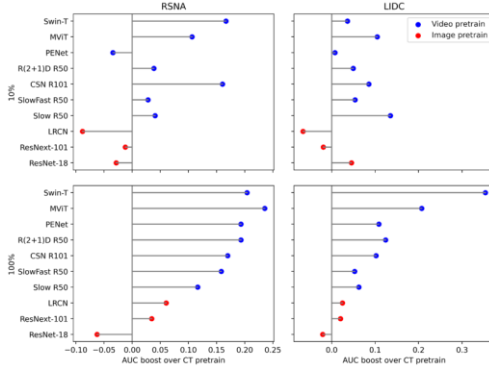


Figure 4: Video pretraining outperforms in-domain CT pretraining