

Sex Detection Based Self-Supervised Learning for Prediction of LVEF from the ECG

James Brundage, Brian Zenger, Jake Bergquist, Ann Lyons, Ryan Butcher, Bao Wang, Rashmee Shah, Rob S. MacLeod, Benjamin Steinberg, Tolga Tasdizen

Supervised deep learning (DL) has become an increasingly common tool for advanced analysis of the electrocardiogram (ECG). These methods rely heavily on labeled datasets, in which there is a clinical annotation for each ECG. However, real world ECG datasets may not contain enough labeled recordings to facilitate robust feature extraction, preventing DL analysis for clinical problems with small datasets. Non-contrastive self-supervised learning (NCSSL) seeks to utilize cheaply labeled data to improve performance in a supervised learning task. This process consists of first training a model on a primary task with cheap data labels, followed by a second training process which attempts to learn the downstream task by initializing with weights learned from the first. This second model can either be fine-tuned, where only the final model layer is allowed to be updated, or initialized, where all model layers are allowed to be updated. Here, we tested the effect of a sex detection based pre-training task on the amount of data needed to train a model to detect low left-ventricular ejection fraction (LVEF). We found that although performance dropped as training set size decreased, NCSSL with a sex detection task ameliorated the decline and surpassed baseline performance with only 50% of training set data. When using 100% of the training set data and sex detection based initialization, the best model performance jumped from an AUC of 0.90 to 0.97 on average. The initialization approach outperformed the fine-tuning approach when using weights from the sex detection task. These findings indicate that NCSSL using sex detection can help leverage unlabeled data to achieve high performance on complex ECG detection tasks.

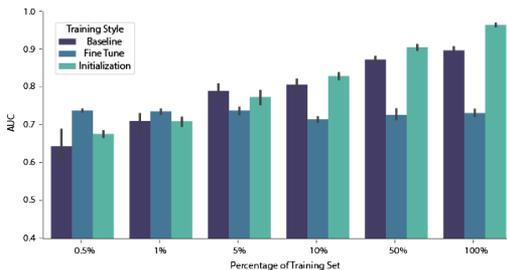


Figure 1. Visualization of the comparison of different sex detection based training approaches by training set size. Values represent mean testing set AUC from 5 different trainings. Error bars represent standard deviation of the mean.