Figure 1: **Training curves using prior models with varying accuracies.** The dashed line represents the training curve of BYOL (baseline), while the solid lines correspond to prior models with different accuracies. "Prior model (BYOL)" indicates the use of BYOL as the prior model, and the accuracies of the other prior models are all lower than that of BYOL. Notably, most prior models surpass BYOL performance in the early stages of training, and even a model with approximately 75% accuracy can outperform BYOL in the final performance. The results demonstrate that *employing a strong model as the prior model is not essential*.