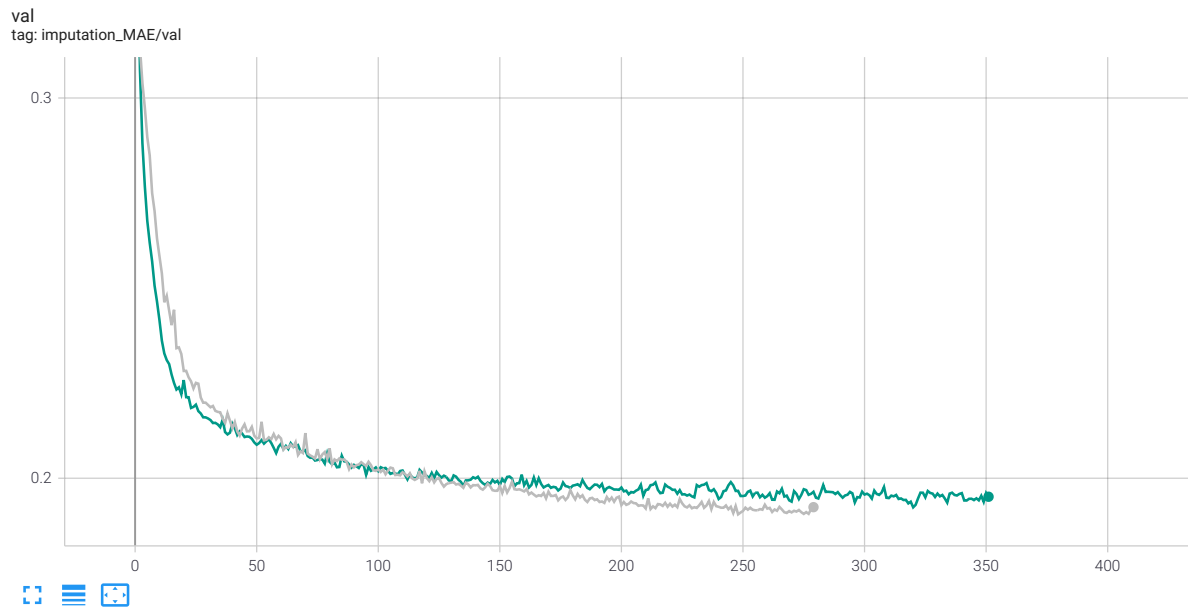


(a) Reconstruction MAE



(b) Imputation MAE

Figure 1: Imputation MAE and reconstruction MAE of SAITS trained on only MIT (the green one) and trained on MIT+ORT (the gray one) in the validation stage. Note that SAITS models used here are both SAITS (base) and both are early stopped according to our training strategy. In Figure 1(a), SAITS (MIT)'s reconstruction loss firstly falls and then rises because SAITS (MIT) is not trained on the reconstruction task. By contrast, SAITS (MIT+ORT) gets well-trained on this task. In Figure 1(b), without ORT, the imputation MAE of SAITS (MIT) can drop slightly faster than SAITS (MIT+ORT)'s, but its final result is a bit worse than SAITS (MIT+ORT)'s. We can think SAITS (MIT) here is a bit overfitting on its only task, namely MIT. We know that imputed values are related to observed data, and ORT makes SAITS (MIT+ORT) have to converge on observed data as well. Therefore, ORT can help the model further optimize on the MIT, though not much.