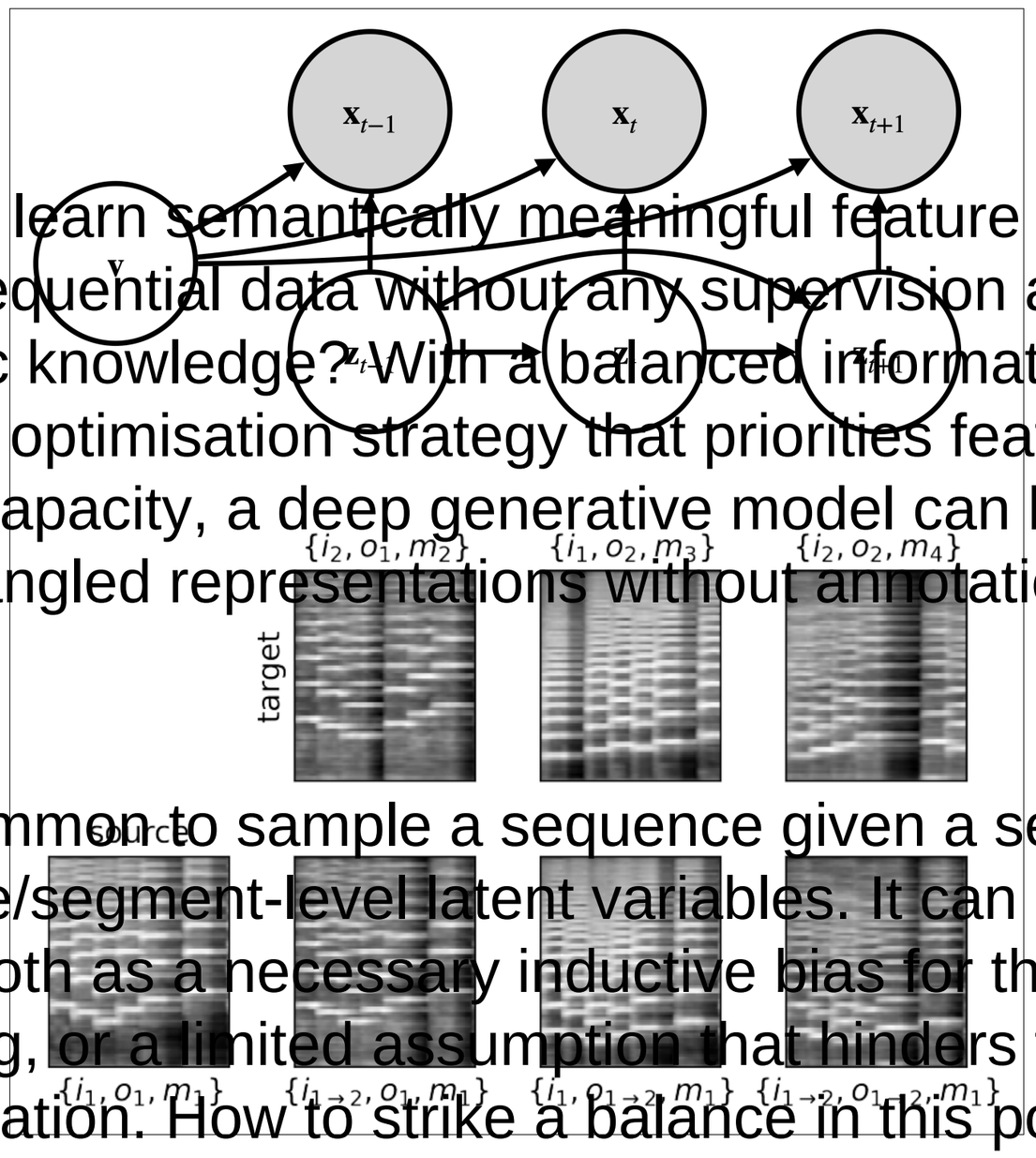
Yin-Jyun Luo

Sebastian Ewert and Simon Dixon

Unsupervised Disentangled Representation Learning for Sequential Data



How to learn semantically meaningful feature representations from sequential data without any supervision and domain-specific knowledge? With a balanced information bottleneck and an optimisation strategy that priorities features of lower capacity, a deep generative model can learn disentangled representations without annotations.

It is common to sample a sequence given a sequence-level and a frame/segment-level latent variables. It can probably be seen both as a necessary inductive bias for the unsupervised learning, or a limited assumption that hinders the optimisation. How to strike a balance in this potential trade-off?