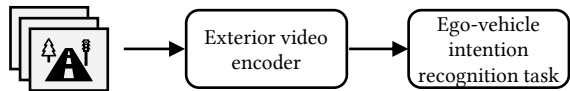
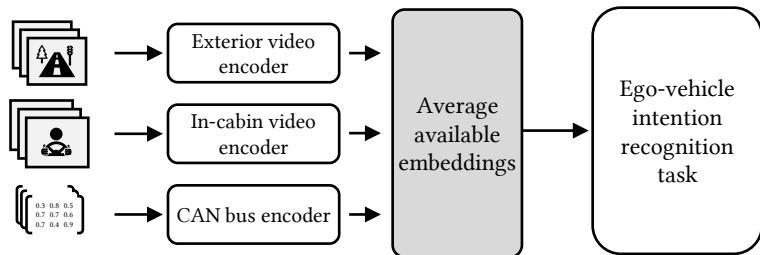


Uni-modal



Multi-modal



Multi-modal with shared encoder

