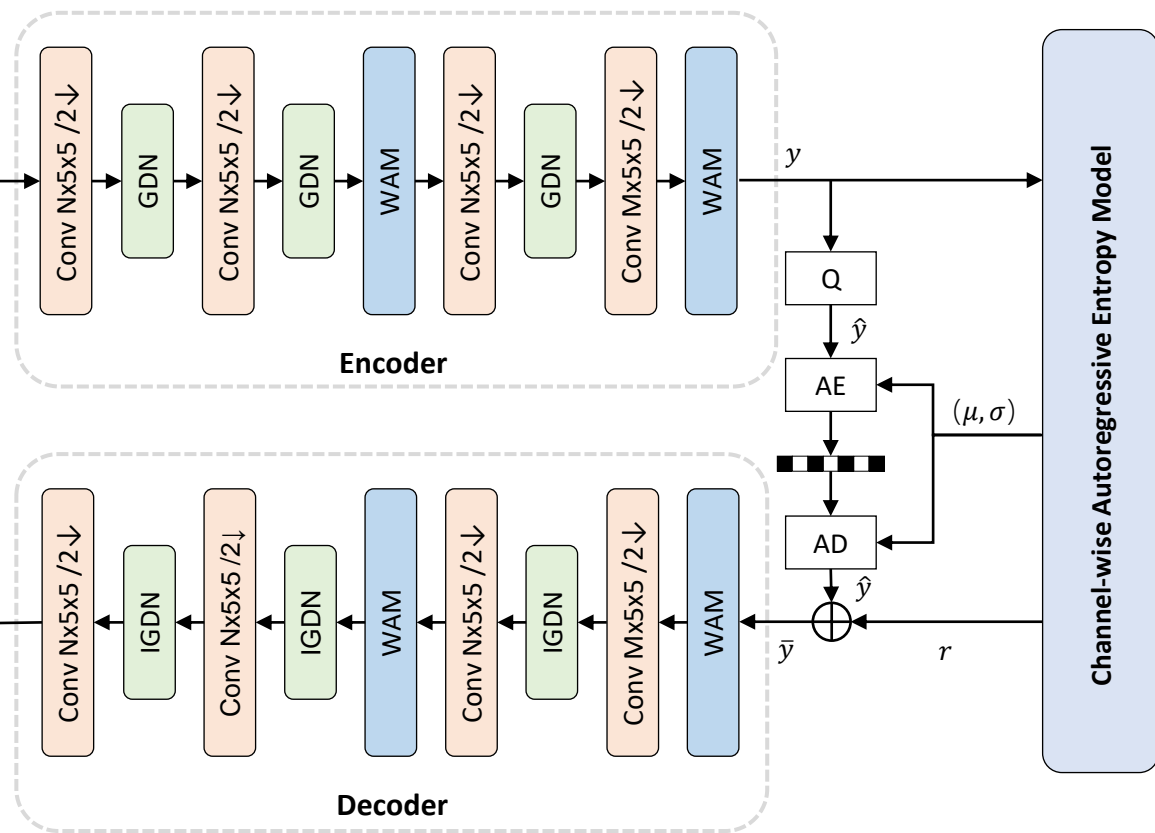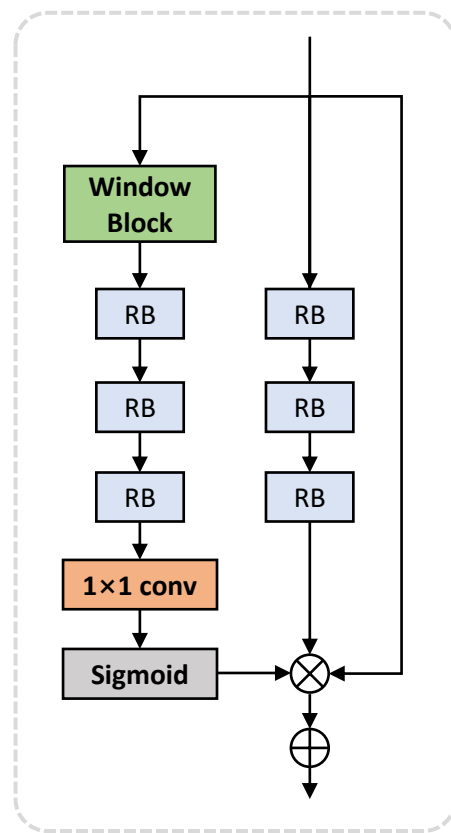(a) The Architecture of Our CNN-based Model

(b) Window Attention Module