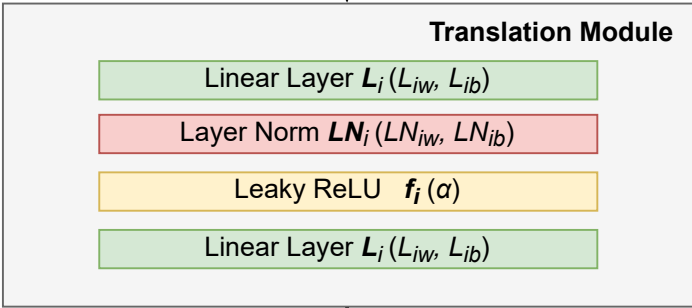


Image
in
Domain **A**



CLIP Full Sequential Embedding

Weighted Average (1D Conv Layer) $\mathbf{C}(W_{Aw}, W_{Ab})$



Layer Norm

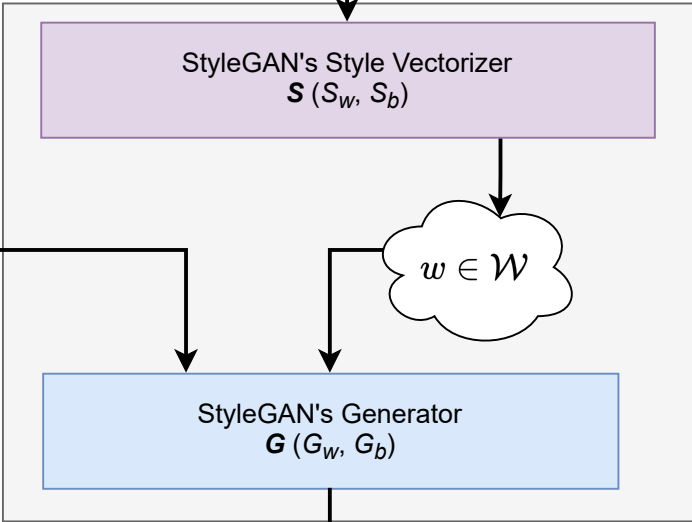


Image
Noise
 $N = \mathcal{G}(0, 1)$

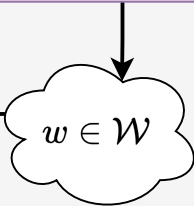


Image
in
Domain **B**