



TE - Target Embedding
 IE - Input Embedding
 MHSA - Multi Head Self Attention
 LSTM - Long Short Term Memory

→ Hidden State Information transfer
 a - Adapter Function