Output task-specific Replaceable Decoder encoded sequence Transformer Encoder Embedding per-word Input word sequence