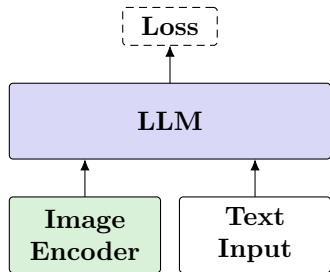


Encoder-Decoder



Adapted LLM (Like Flamingo)