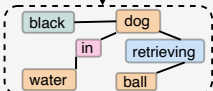


Source: A black dog is retrieving a ball in water



Stable Diffusion



Target: Ein schwarzer Hund holt einen Ball aus dem Wasser.



+



Vision Encoder+ MLP

LLM Embedding



Imagine

Source: A black dog is retrieving a ball in water