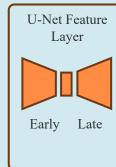


-- DvVA-Owen2.5-7B



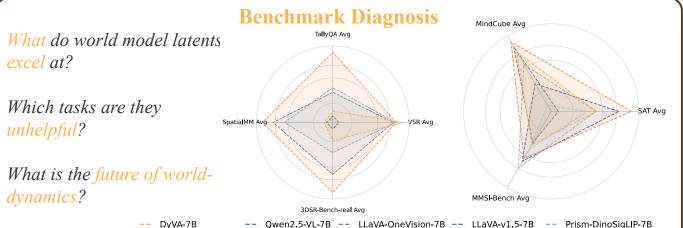
**Design Exploration** 

*features do we need?* What kind of semantics might complement these features?

Can Standalone Video Generation Features work?

-- Prism-SigLIP-7B

How much world model



Intern2.5-VL-8B