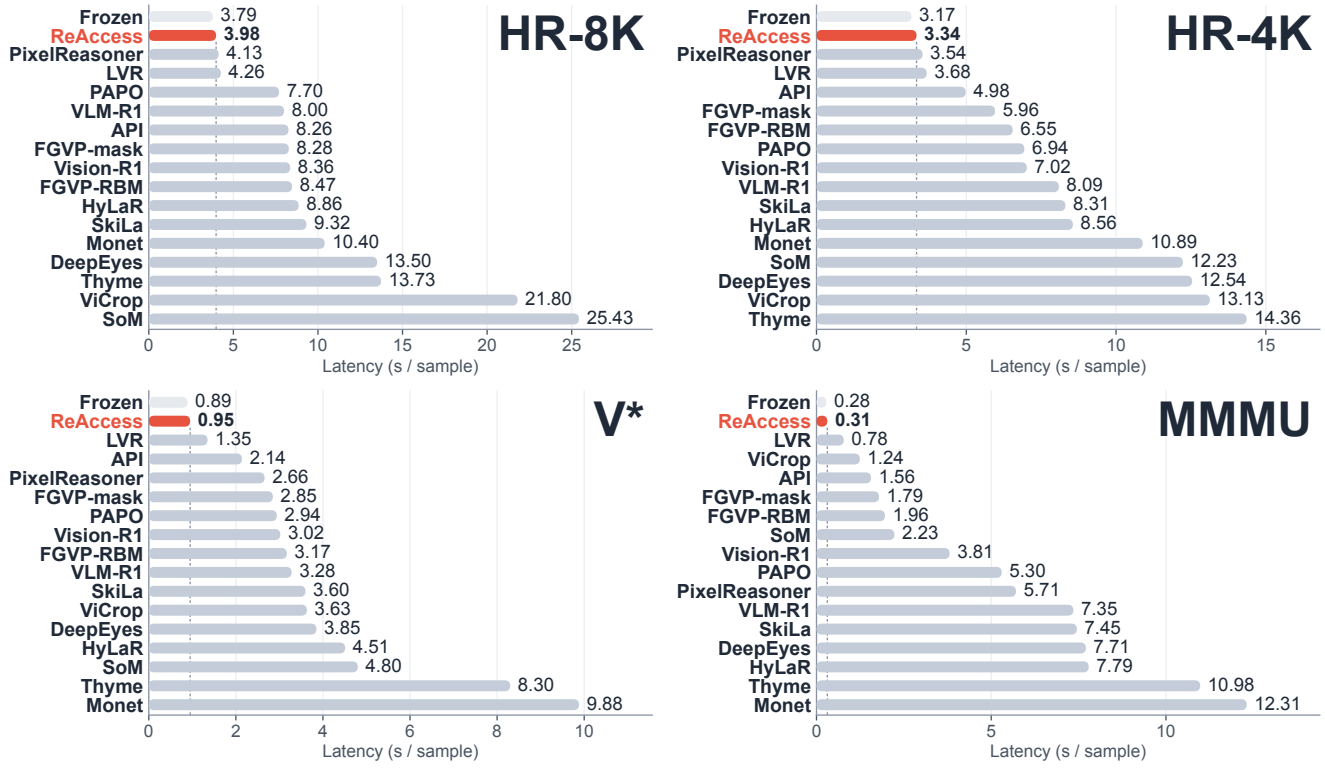

HR-8K
Frozen 3.79
ReAccess 3.98
PixelReasoner 4.13
LVR 4.26
PAPO 7.70
VLM-R1 8.00
API 8.26
FGVP-mask 8.28
Vision-R1 8.36
FGVP-RBM 8.47
HyLaR 8.86
SkiLa 9.32
Monet 10.40
DeepEyes 13.50
Thyme 13.73
ViCrop 21.80
SoM 25.43
0 5 10 15 20 25
Latency (s / sample)
HR-4K
Frozen 3.17
ReAccess 3.34
PixelReasoner 3.54
LVR 3.68
API 4.98
FGVP-mask 5.96
FGVP-RBM 6.55
PAPO 6.94
Vision-R1 7.02
VLM-R1 8.09
SkiLa 8.31
HyLaR 8.56
Monet 10.89
SoM 12.23
DeepEyes 12.54
ViCrop 13.13
Thyme 14.36
0 5 10 15
Latency (s / sample)
V*
Frozen 0.89
ReAccess 0.95
LVR 1.35
API 2.14
PixelReasoner 2.66
FGVP-mask 2.85
PAPO 2.94
Vision-R1 3.02
FGVP-RBM 3.17
VLM-R1 3.28
SkiLa 3.60
ViCrop 3.63
DeepEyes 3.85
HyLaR 4.51
SoM 4.80
Thyme 8.30
Monet 9.88
0 2 4 6 8 10
Latency (s / sample)
MMMU
Frozen 0.28
ReAccess 0.31
LVR 0.78
ViCrop 1.24
API 1.56
FGVP-mask 1.79
FGVP-RBM 1.96
SoM 2.23
Vision-R1 3.81
PAPO 5.30
PixelReasoner 5.71
VLM-R1 7.35
SkiLa 7.45
DeepEyes 7.71
HyLaR 7.79
Thyme 10.98
Monet 12.31
0 5 10
Latency (s / sample)