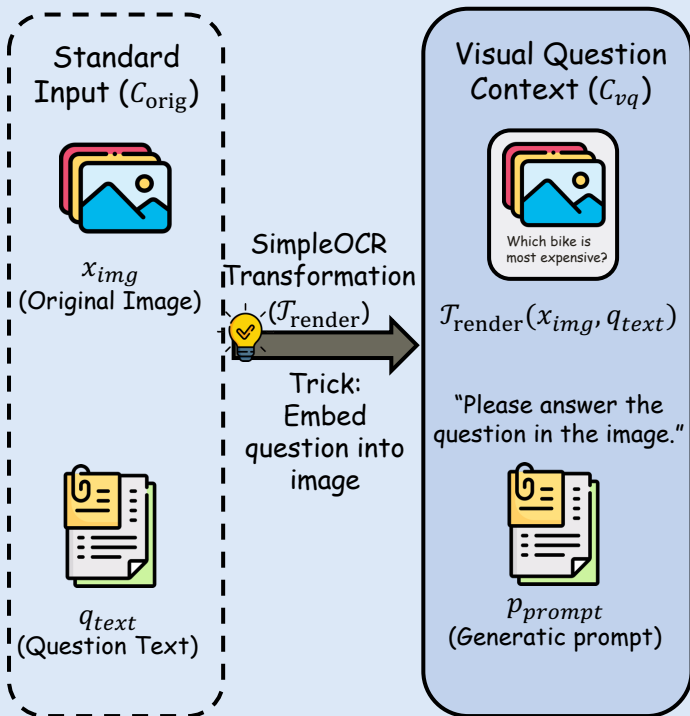
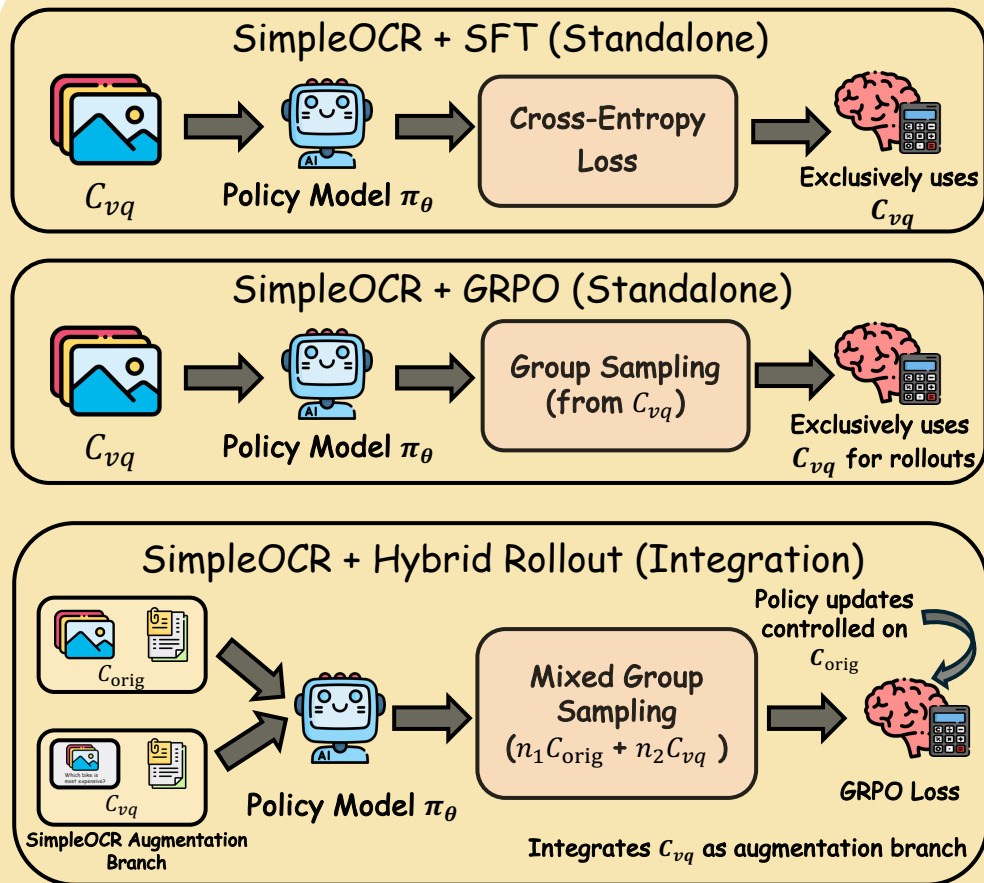


a) Data Transformation (SimpleOCR)



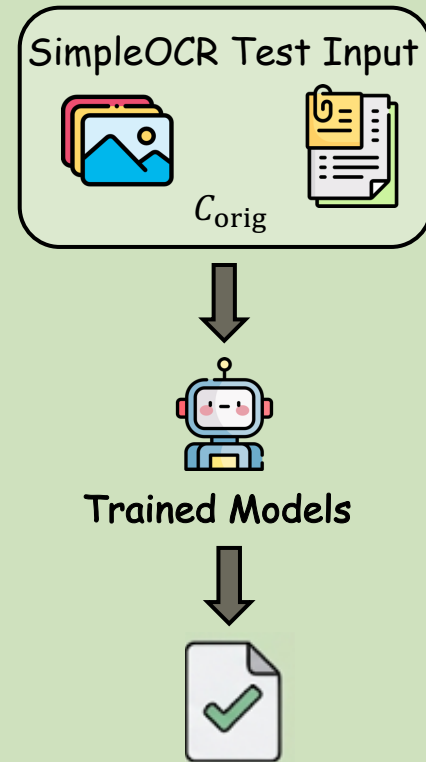
Create visual-only context

b) Training Phase (Forcing Visual Engagement)



SimpleOCR(C_{vq}) forces visual engagement either as sole input or play and plug

c) Inference / Evaluation



Evaluate on standard format