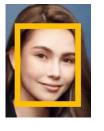


(1) Box

Identify and locate faces in the image.



(2) Box + Number **Enhance grounding capabilities** by box and number





Left Eye

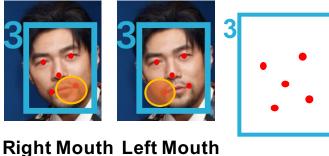
Right Eye



Nose







(3) Box + Number + Facial Landmarks **Analyze facial expression by spatial** relationships

