



(a) Query patch in input image

(b) Instance-dependent connectivity mask

(c) Sparse masked attention heatmap

(d) Full attention heatmap