6813.jpg				
	The same of the same	The same of the same	AND DESCRIPTIONS	87.70146942138672

7422.jpg

83.20631408691406

Figure 2: Visualization results of attention weights on each stage and predicted density maps from MixformerCAC. References are annotated with bounding boxes, and we use red color to indicate the reference selected to visualize attention weights.