

Representing procedure text such as recipe for crossmodal retrieval is inherently a difficult problem, not mentioning to generate image from recipe for visualization