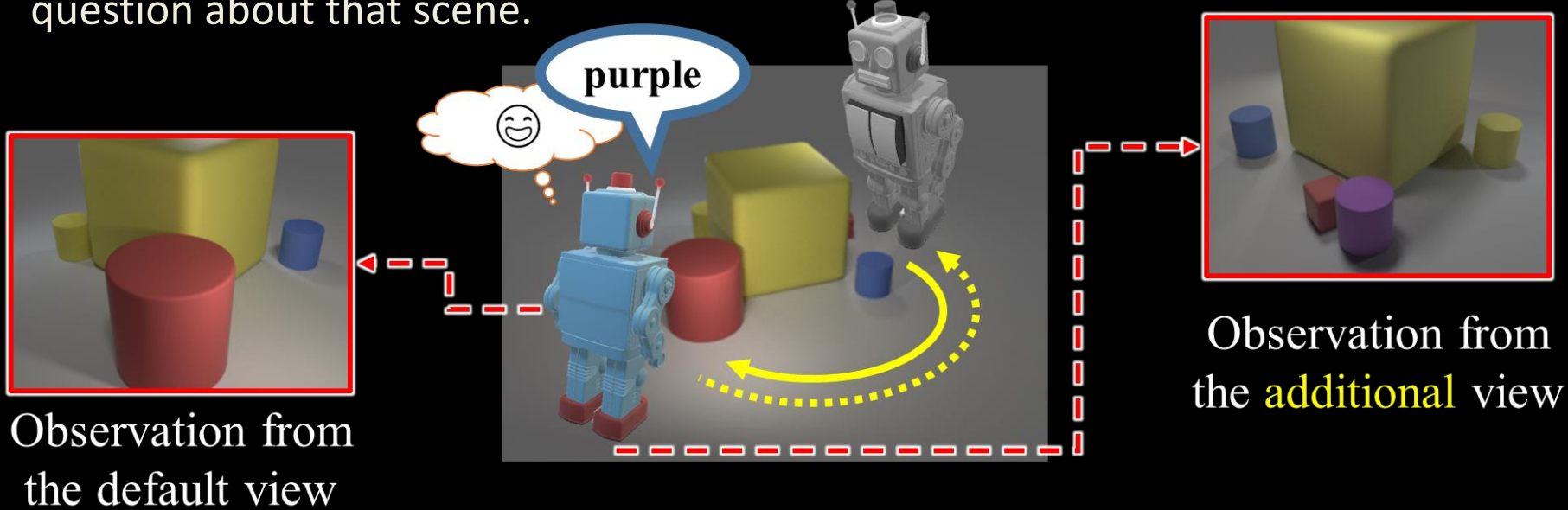


# Multi-View Visual Question Answering with Active Viewpoint Selection

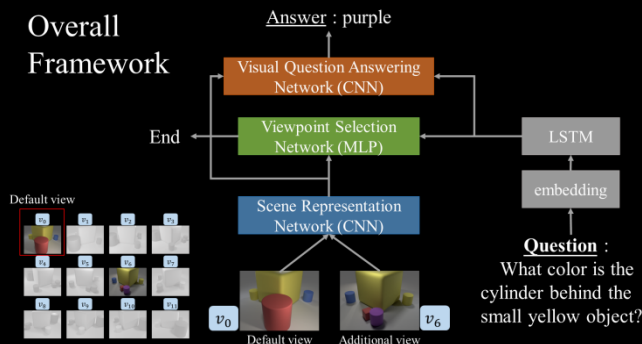
Yue Qiu <sup>1,2</sup>, Yutaka Satoh <sup>2,1</sup>, Ryota Suzuki <sup>2</sup>, Kenji Iwata <sup>2</sup>, Hirokatsu Kataoka <sup>2</sup>  
<sup>1</sup>University of Tsukuba, <sup>2</sup>National Institute of Advanced Industrial Science and Technology

- A framework that observes the scene iteratively in order to answer a given question about that scene.



**Question** : What color is the cylinder behind the small yellow object?

Implementing the method by constructing a new deep neural network architecture including **VQA**, **Viewpoint Selection** and **Scene Representation** modules.



NATIONAL INSTITUTE OF  
ADVANCED INDUSTRIAL SCIENCE  
AND TECHNOLOGY (AIST)