

We propose a novel model to address the task of Visual Dialog which exhibits complex dialog structures. To obtain a reasonable answer based on the current q