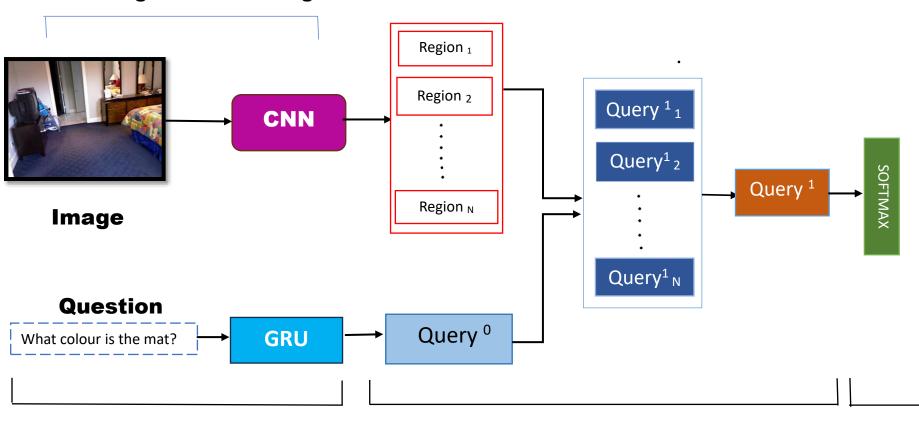
Image Understanding



Question Encoding

Reasoning

Answering