

Our goal in this work is to train an image captioning model that generates more dense and informative captions. We introduce "relational captioning," a novel im