Though image-to-sequence generation models have become overwhelmingly popular in human-computer communications and the sequence of the sequenc	nications, they suffer from strongly favoring sa