

We propose to compose dynamic tree structures that place the objects in an image into a visual context, helping visual reasoning tasks such as scene graph ge