

To understand a scene in depth not only involves locating/recognizing individual objects, but also requires to infer the relationships and interactions among them