Common object counting in a natural scene is a challenging problem in computer vision with numerous real-world applications. Existing image-level supervised