Conventional zero-shot learning (ZSL) methods generally learn an embedding, e.g., visual-semantic mapping, to handle the unseen visual samples via an indire