

Learning good feature embeddings for images often requires substantial training data. As a consequence, in settings where training data is limited (e.g., few-sho