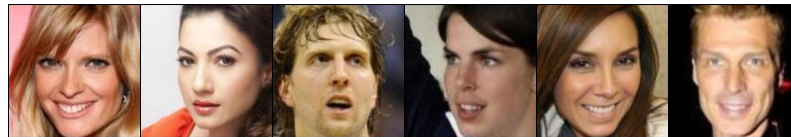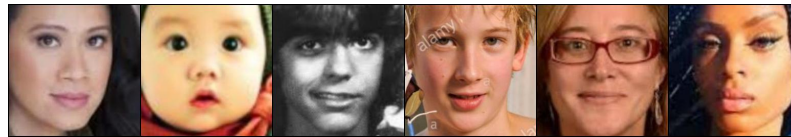# A. More Results

(a) (**CIFAR-10**, STL-10)
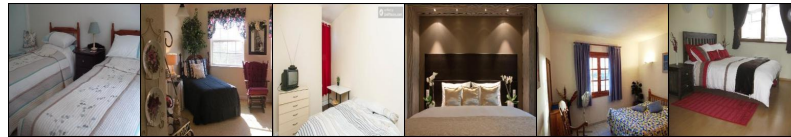

(b) (CIFAR-10, **STL-10**)

(c) (**CelebA**, UTKFace)
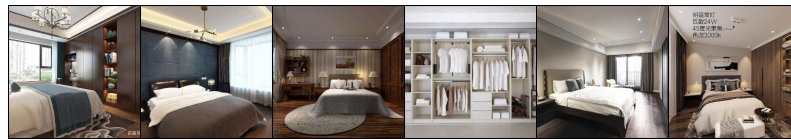

(d) (CelebA, **UTKFace**)

(e) (**LSUN-Bed**, Wild Bed)

(f) (LSUN-Bed, **Wild Bed**)

(g) (**LSUN-Church**, Wild Church)

(h) (LSUN-Church, **Wild Church**)

Figure 7. The complete version of image examples of different (member, auxiliary) dataset pairs, where the images are sampled from the **boldface dataset**, and "Bed" is short for "Bedroom".
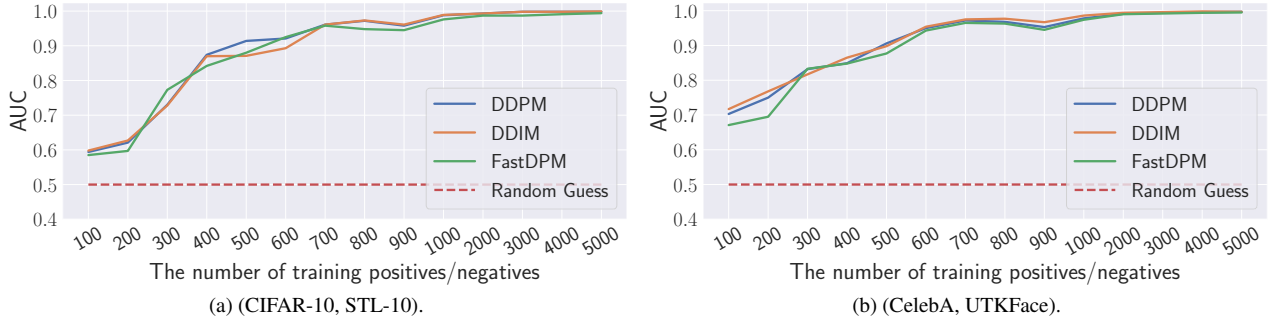
(a) (CIFAR-10, STL-10).

(b) (CelebA, UTKFace).

Figure 8. The complete version of attack performances with different query budgets in unconditional generation tasks.

(a) Unconditional generation tasks.
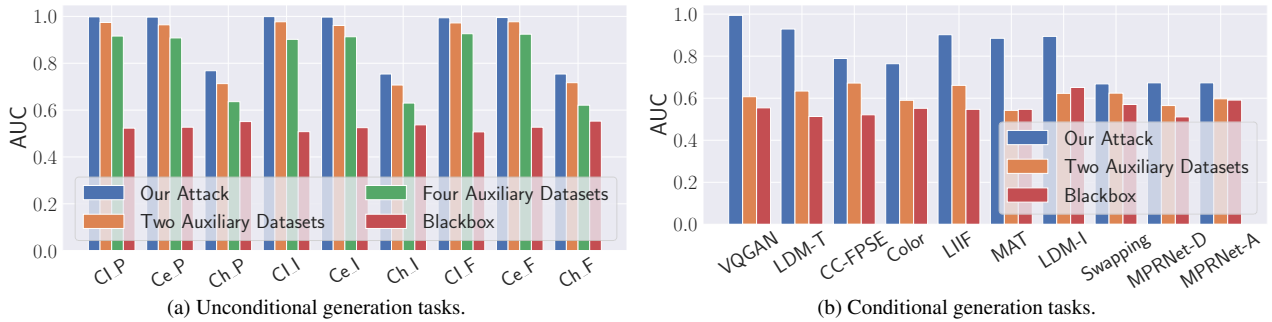

(b) Conditional generation tasks.

Figure 9. The complete version of attack performances with different numbers of auxiliary datasets.

B. Datasets

In this section, we will introduce the datasets used in our experiments, including CIFAR-10, STL-10, CelebA, UTKFace, LSUN, ImageNet, Open Images, LAION, COCO, ADE20K, and SIDD.

- **CIFAR-10.** The CIFAR-10 dataset consists of $60,000$ images of 10 classes, i.e., airplane, automobile, bird, cat, deer, dog, frog, horse, ship, and truck. For each class, there are $5,000$ training images and $1,000$ test images.

- **STL-10.** The STL-10 dataset is inspired by the CIFAR-10 dataset but with some modifications, covering the classes of airplane, bird, car, cat, deer, dog, horse, monkey, ship, and truck. And for each class, there are 500 training images and 800 test images. Besides, there are $100,000$ unlabeled images for unsupervised learning.

- **CelebA.** The CelebA dataset contains more than $200,000$ celebrity images with different face attributes. Each face image has 40 attribute annotations with binary values, indicating whether this image satisfies corresponding attributes, e.g., whether wearing a hat.

- **UTKFace.** The UTKFace dataset contains $20,000$ face images with annotations of age, gender, and ethnicity, with a long age span (range from 0 to 116 years old).

- **LSUN.** The LSUN dataset contains 10 scene categories, e.g., bedroom and church. For LSUN-Bedroom, there are more than $3,000,000$ training images and 300 validation images. And for LSUN-Church, there are more than $120,000$ training images and 300 validation images.

- **ImageNet.** The ImageNet dataset contains more than $14,000,000$ annotated images according to the WorldNet hierarchy. In the paper, we use its tiny version, i.e., Tiny ImageNet. The Tiny ImageNet dataset covers 200 classes, and each class has 500 training images, 50 test images, and 50 validation images.

- **Open Images.** The Open Images dataset consists of 9 million training images, which are partially annotated, with $9,600$ trainable classes.

- **LAION.** The LAION-400M dataset contains 400 million English (image, text) pairs. And all images and texts are filtered by CLIP [42] according to the cosine similarity between the image and text embeddings.

- **COCO.** The COCO dataset is a large-scale dataset containing $330,000$ images, within which more than $200,000$ images are labeled with 80 object categories.

- **ADE20K.** The ADE20K dataset contains more than $20,000$ scene-centric images which are annotated with pixel-level objects and object parts labels. There are totally 150 semantic categories, including stuffs and discrete objects.

- **SIDD.** The SIDD dataset consists of $30,000$ noisy images taken by 5 representative smartphone cameras from 10 scenes under different lighting conditions.

C. Training Details

For label balance, the number of training positives is equal to that of training negative in our experiments. If there is no limitation on query times, it is possible to generate any desired number of training positives. Thus, the number of training samples relies on the size of auxiliary datasets (where the training negatives are derived). In our experiments, in any case that the auxiliary dataset could provide more than $5,000$ training negatives, we will set the number of training samples to $10,000$ (i.e., $5,000$ for training positives and $5,000$ for training negatives).