(a). Visual examples of the proposed DF40 dataset.

Known Domain

Face Swap (FS).
Total number of FS: 10

MobileSwap   SimSwap   BlendFace   e4s

Face Reenact (FR)
Total number of FR: 13

SadTalker   FOMM   Lia   Wav2Lip

Entire Face Synthesis (EFS)
Total number of EFS: 12

VQGAN   StyleGAN-XL   SD-2.1   DiT

Unknown Domain

Face Editing (FE)
Total number of FE: 5

StyleCLIP   StarGANv2   CollabDiff   e4e

Remaining open-set Synthesis deepfakes
Total number: 4

HeyGen   DeepFaceLab   MidJourney-6   WhichIsReal

Statistic Details of DF40:

Known Domains:
1. Num of FS: 9
2. Num of FR: 12
3. Num of EFS: 10
Total Num: 31

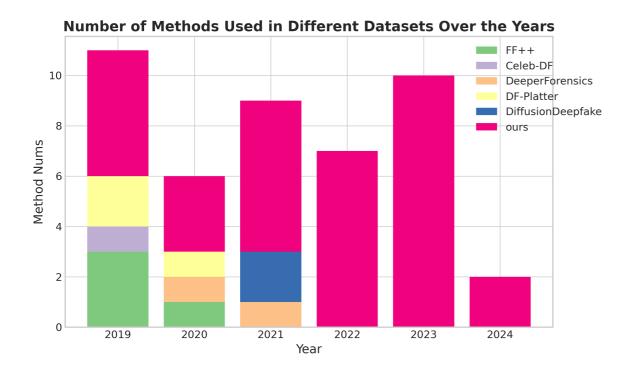*Unknown Domains:*
1. Num of FS: 1;
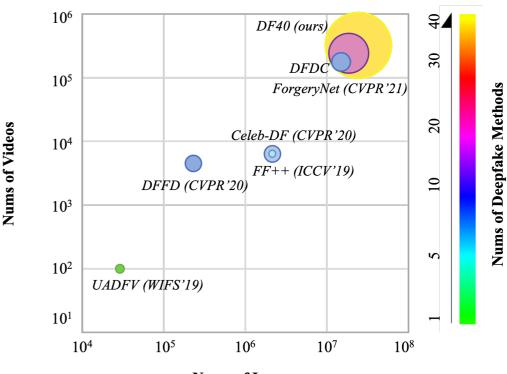2. Num of FR: 1;
3. Num of EFS: 2;
4. Num of FE: 5
Total Num: 9

Note: *Known domains* refer to real data from *FF++* and *CDF,* while ***unknown domains*** are those from other real datasets

(b). Statistic comparison of forgery number used in different datasets

(c). Comparison with other dataset in *diversity* and scale.