PRESENTATION ON
DEEPPRIVACY: A GAN FOR FACE ANONYMIZATION
(ARXIV: 1909.04538)

Pratikkumar.Prajapati@sjsu.edu

Apr/07/2020

OVERVIEW

- A novel architecture which is able to automatically anonymize faces in images
- Anonymizes faces without destroying the existing data distribution, i.e. face realistically fits the given situation in the image.
- Based on a conditional generative adversarial network (cGAN)
- Model trained for 17 days on two NVIDIA V100-32GB GPU on 40M images

SAMPLE RESULTS

Fig. 1: **DeepPrivacy Results** on a diverse set of images. The left image is the original image annotated with bounding box and keypoints, the middle image is the input image to our GAN, and the right image is the generated image. Note that our generator never sees any privacy-sensitive information.

DATASET

- *Flickr Diverse Faces* (FDF)
  - A new dataset provided by the authors, crawled from the YFCC-100M dataset
  - Used for training only
  - consists of 1.47M faces, with bounding box and keypoint annotation for each face.
- validation performed on the WIDER-Face dataset

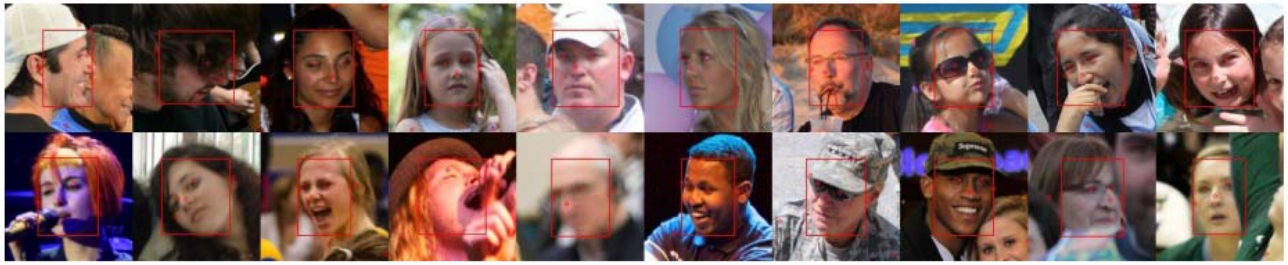

Fig. 2: **The FDF dataset**. Each image has a sparse keypoint annotation (7 keypoints) of the face and a tight bounding box annotation. We recommend the reader to zoom in.

DATA PRE-PROCESSING

- Two simple annotations of the face
  1. A bounding box annotation to identify the privacy-sensitive area, e.g. a face
     - Used Single Shot Scale invariant Face Detector for bounding box
  2. A sparse pose estimation of the face, containing keypoints of the face
     - 7 Keypoints: left/right eye, left/right ear, left/right shoulder, and nose.
     - Used mask R-CNN, with a ResNet-50 FPN backbone for keypoint estimation
     - one-hot encoded image of size $K \times M \times M$, where K is the number of keypoints and M is the
       target resolution.

DEEPPRIVACY MODEL

- Based on cGAN
- Progressive growing training technique
  - For both Generator and Discriminator
  - From starting resolution of 8×8 to 128×128
- Considers the existing background and a sparse pose annotation to generate realistic anonymized faces.
- Generator never observes the original face

DEEPPRIVACY - GENERATOR

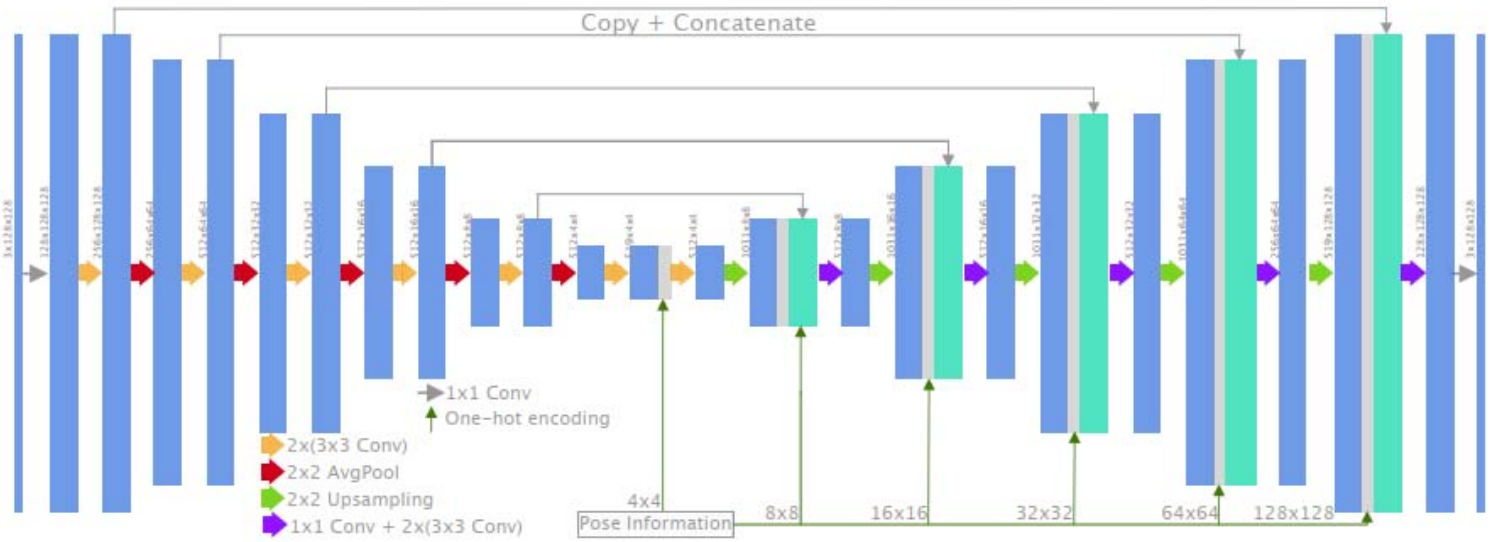

Fig. 3: **Generator Architecture** for 128×128 resolution. Each convolutional layer is followed by pixel normalization [12] and LeakyReLU($\alpha = 0.2$). After each upsampling layer, we concatenate the upsampled output with pose information and the corresponding skip connection.

DEEPPRIVACY - DISCRIMINATOR

- Includes the background information as conditional input to the start of the discriminator, making the input image have six channels instead of three.
- Pose information is concatenated with the output of each down sampling layer
- Deep discriminator
  - doubles the number of convolutional layers for each resolution.
  - To mimic the skip-connections in the generator, we wrap the convolutions for each resolution in residual blocks.
- Wide discriminator
  - Same architecture; however, we increase the number of filters in each convolutional layer by a factor of $\sqrt{2}$

SAMPLE OUTPUT

- Converges to a Frechect Inception Distance (FID) of 1.53

Fig. 4: **Anonymized Images from DeepPrivacy**. Every single face in the images has been generated. We recommend the reader to zoom in.

RESULTS (1/2)

Table 1: **Face Detection AP** on the WIDER Face [27] validation dataset. The face detection method used is DSFD [14], the current state-of-the-art on WIDER-Face.

Anonymization method	Easy	Medium	Hard
No Anonymization [14]	96.6%	95.7%	90.4%
Blacked out	24.9%	36.3%	54.8%
Pixelation ($16x16$)	95.3%	94.9%	**90.2%**
Pixelation ($8x8$)	91.4%	92.3%	88.9%
9x9 Gaussian Blur ($\sigma = 3$)	95.3%	92.8%	84.7%
Heavy Blur (filter size = 30% face width)	83.4%	86.3%	86.1%
DeepPrivacy (Ours)	**95.9%**	**95.0%**	89.8%

Fig. 5: **Different Anonymization Methods** on a face in the WIDER Face validation set.

RESULTS (2/2)

Table 2: **Ablation Experiments** with our model. We report the Frèchet Inception Distance (FID) on the FDF validation dataset, after showing the discriminator $30.0M$ images (lower is better). For results in Table 2a and Table 2b, we use a model size of $12M$ parameters for both the generator and discriminator. *Reported after $20.0M$ images, as the deep discriminator diverged after this.

(a) Result of using conditional pose.

Model	FID
With Pose	**2.71**
Without Pose	3.36

(b) Result of the deep and wide discriminator.

Discriminator	FID
Deep Discriminator*	9.327
Wide Discriminator*	**3.86**

(c) Result of different model sizes.

#parameters	FID
12M	2.71
46M	**1.84**

LIMITATION

Fig. 6: **Failure Cases of DeepPrivacy** Our proposed solution can generate unrealistic images in cases of high occlusion, difficult background information, and irregular poses.

REFERENCES

- [1] Hukkelas, H.; Mester, R.; and Lindseth, F. 2019. Deepprivacy: A generative adversarial network for face anonymization. In Bebis, G.; Boyle, R.; Parvin, B.; Koracin, D.; Ushizima, D.; Chai, S.; Sueda, S.; Lin, X.; Lu, A.; Thalmann, D.; Wang, C.; and Xu, P., eds., Advances in Visual Computing, 565-578. Cham: Springer International Publishing.