Synthetic Faces

Stephane Gentric Ph.D.
Research Unit Manager
Senior Expert

IDEMIA

IFPC 2020

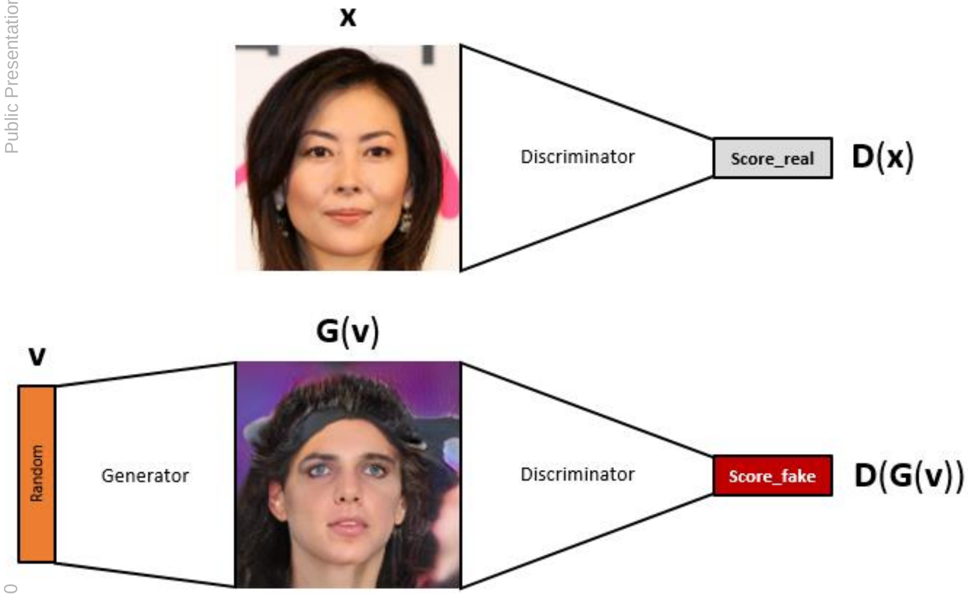
28/10/2020

GAN output: Which are real, which are fake ?

GAN output: Which are real, which are fake ?

Karras, T., Aila, T., Laine, S., & Lehtinen, J. Progressive growing of gans for improved quality, stability, and variation.

What is a GAN ?



“Wasserstein GAN” (2017)

Arjovsky, Chintala, Bottou

To train D minimise:

$$L_D = \frac{1}{N}\sum_{i=1}^{N} \boldsymbol{D}(\boldsymbol{x}_i) - \boldsymbol{D}(\boldsymbol{G}(\boldsymbol{v}_i)) + \text{"}GP\ term\text{"}$$

To train G minimise:

$$L_G = \frac{1}{N}\sum_{i=1}^{N} \boldsymbol{D}(\boldsymbol{G}(\boldsymbol{v}_i))$$

E.g. “Improved Training of Wasserstein GANs” (2017), Gulrajani, ..., Courville

Contents

1. Do GANs generate new identities?
   - Applications: Dataset-anonymisation; semi-supervised learning with distractor images

2. Can GAN Images be used in Biometric Systems?
   - Applications: Non-regression tests; Performances extrapolation; Algorithm-improvement evaluation; Data-augmentation with synthetic sets of mated images …

1. Do GANs generate new identities?

Typical qualitative analysis of overfitting

Visual, nearest-neighbour analyses are typically performed…

Synthetic images:

Nearest neighbours in VGGNet feature-space:

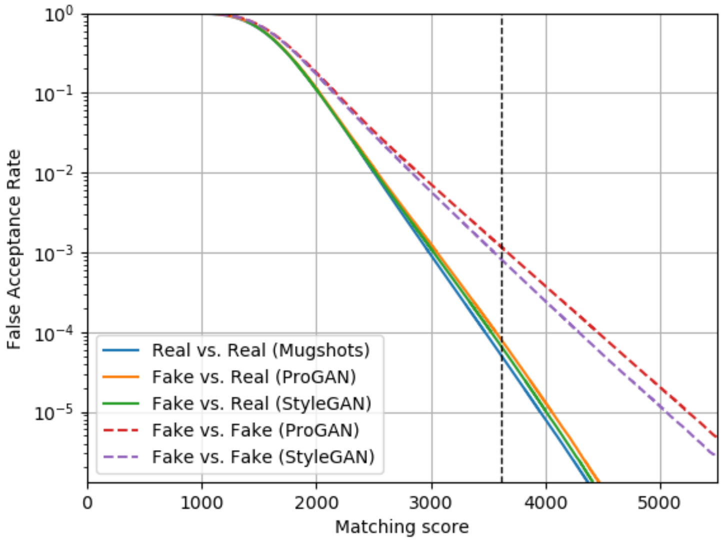
Images taken from *Karras et al. (2018)*

- What about the rest of the space? How frequently do look-alikes occur?

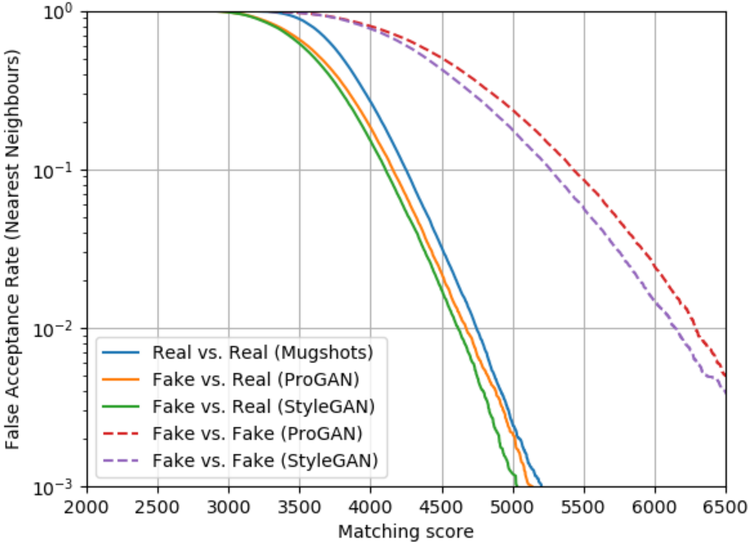
Assessing the frequency of look-alikes (Nearest neighbors)

Matching scores generated using a recent biometric network

All Matching scores



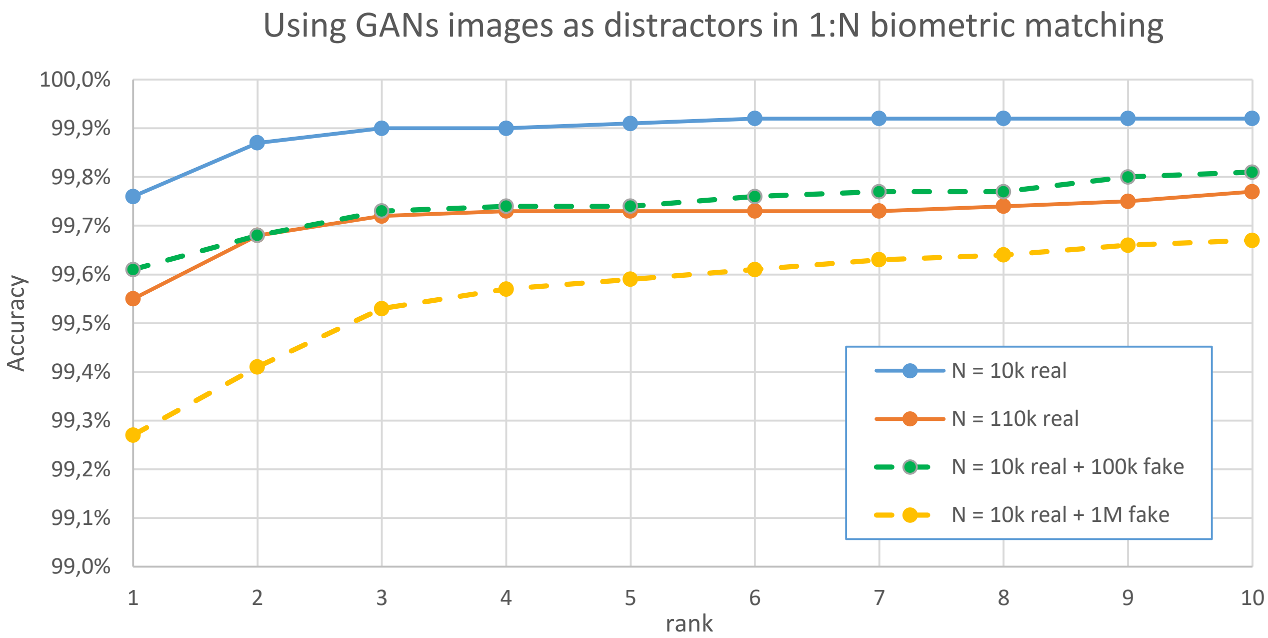
Nearest neighbors scores



ProGAN : Progressive Growing of GANs for Improved Quality, Stability, and Variation. T.Karras & All, ICLR. 2018
StyleGAN : A style-based generator architecture for generative adversarial networks. T.Karras & All, CVPR. 2019

2. Can we use synthetic images as distractors?

Performances with real and fake distractors

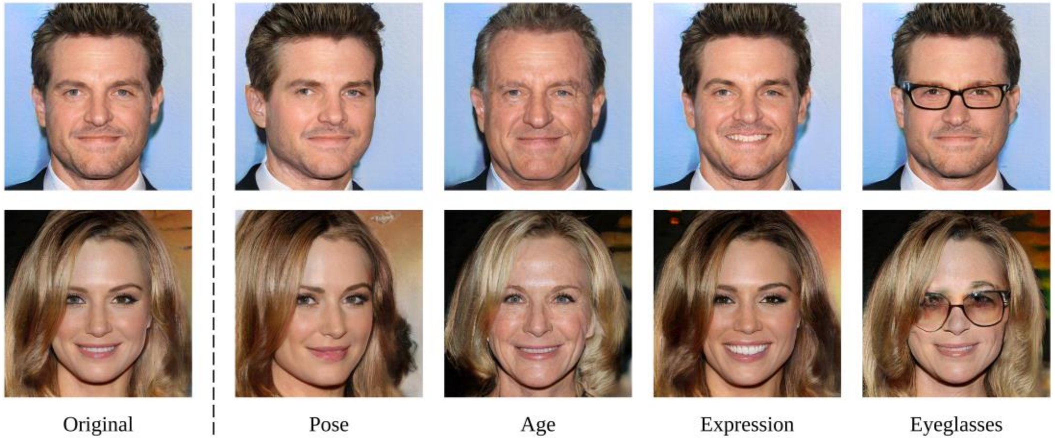


- Synthetic images can be used as distractors to estimate performance with real data.
- As images from GANs exhibit less variability, extrapolation for very large datasets is still uncertain.

3. Can we generate multiple images of the same fake identity?

InterFaceGAN, *Shen et al. (2020)*

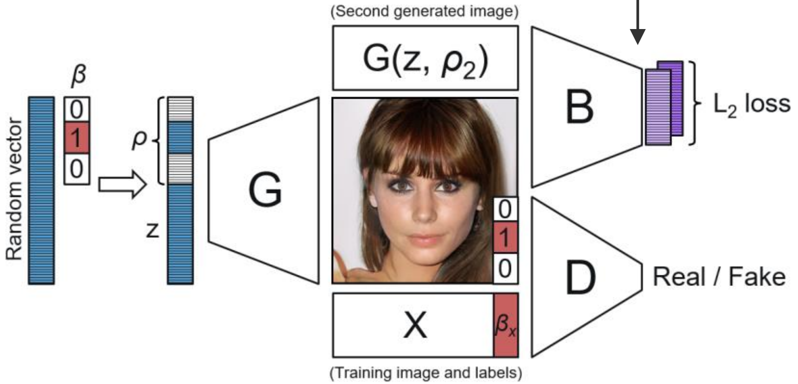
The “InterFaceGAN” method manipulates attributes by traversing the GAN’s latent space in the direction perpendicular to a particular decision boundary.



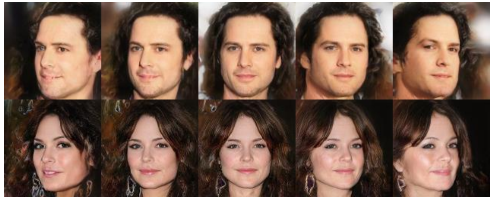
Images taken from *Shen et al. (2020)*

IVI-GAN, *Marriott et.al. (arxiv 2018 – FG 2020)*

Pre-trained
biometric constraint



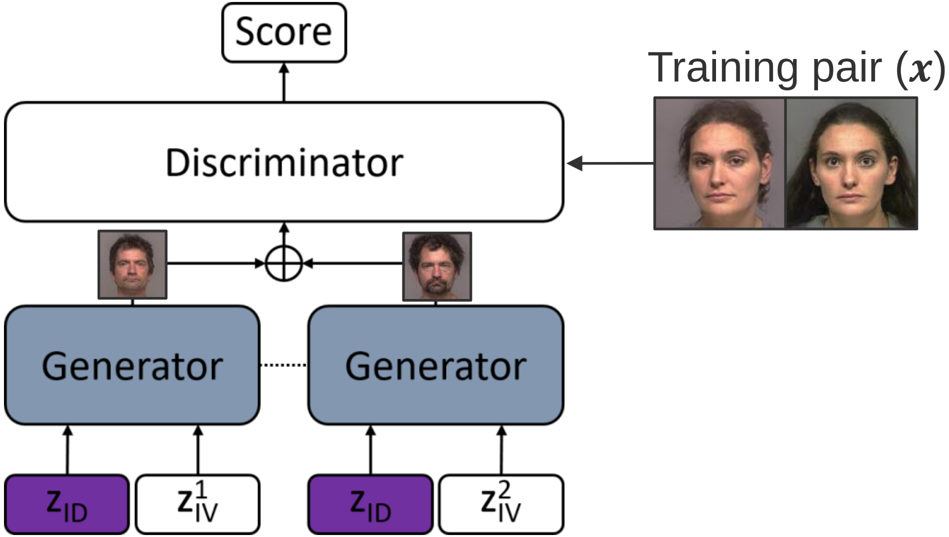
Pose



Background

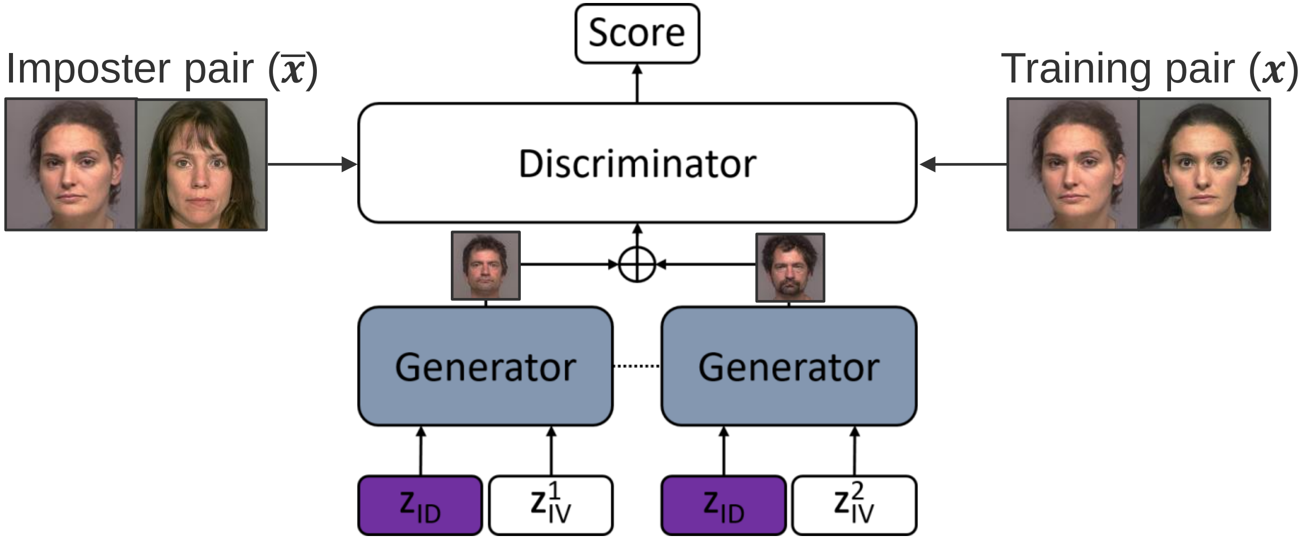
Lighting

SD-GAN, *Donahue et.al. (2018)*



$$\mathcal{L}_{\theta_D} = D(\mathbf{x}) - D(G(\mathbf{z}_1), G(\mathbf{z}_2))$$
$$\mathcal{L}_{\theta_G} = D(G(\mathbf{z}_1), G(\mathbf{z}_2))$$

SD-GAN + Triplet loss, *Marriott et.al. (IJCB 2020)*



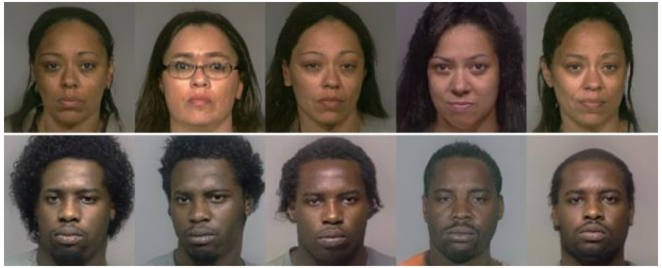
$$\mathcal{L}_{\theta_D} = D(\mathbf{x}) - \frac{1}{2}[D(G(\mathbf{z}_1), G(\mathbf{z}_2)) + D(\bar{\mathbf{x}})] + \lambda[D(G(\mathbf{z}_1), G(\mathbf{z}_2)) - D(\bar{\mathbf{x}})]^2$$

$$\mathcal{L}_{\theta_G} = D(G(\mathbf{z}_1), G(\mathbf{z}_2))$$

Disentangled Datasets

InterFaceGAN (CelebA-HQ)

SD-GAN (Mugshots)



IVI-GAN (CelebA)

SD-GAN + Triplet (Mugshots)

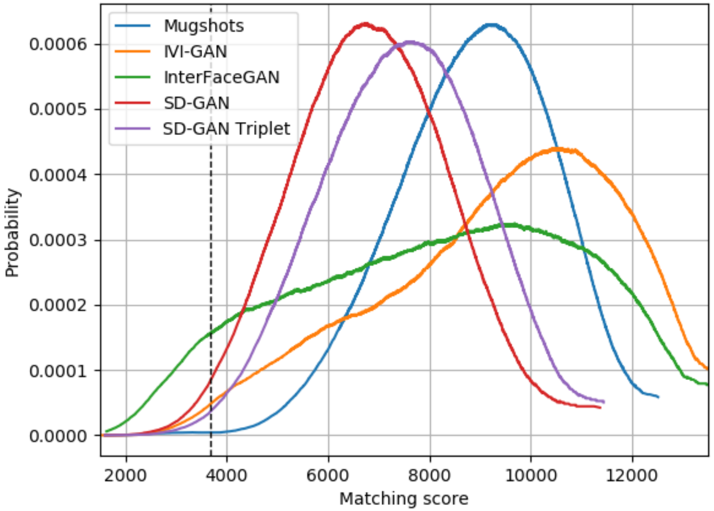
Public Presentation

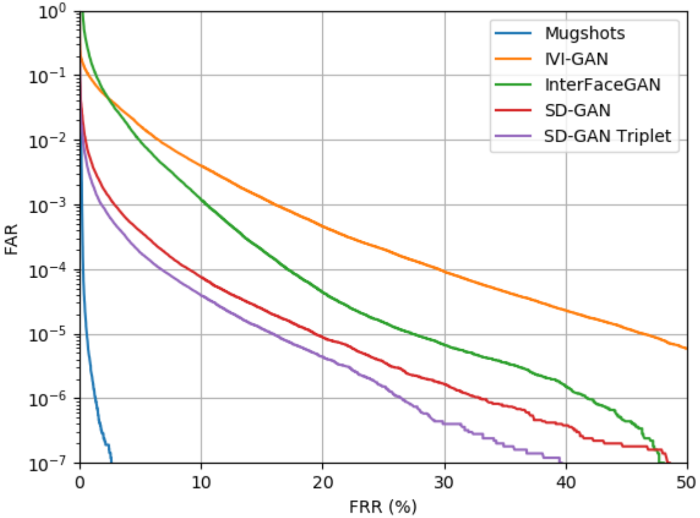
28/10/2020

4. Can GAN Images be used in Biometric Systems?

Biometric scores for mated pairs

Matching score distributions within mated sets, with default parameters



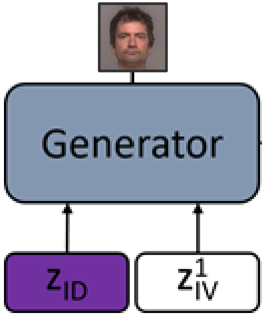


- Mated pairs of synthetic images can be used to compute biometric scores.

- Scores are higher for imposter tests with GAN images.
- Identities are not so well disentangled from other attributes.
- All methods have explicit or implicit parameters leading to different intra-class variability.

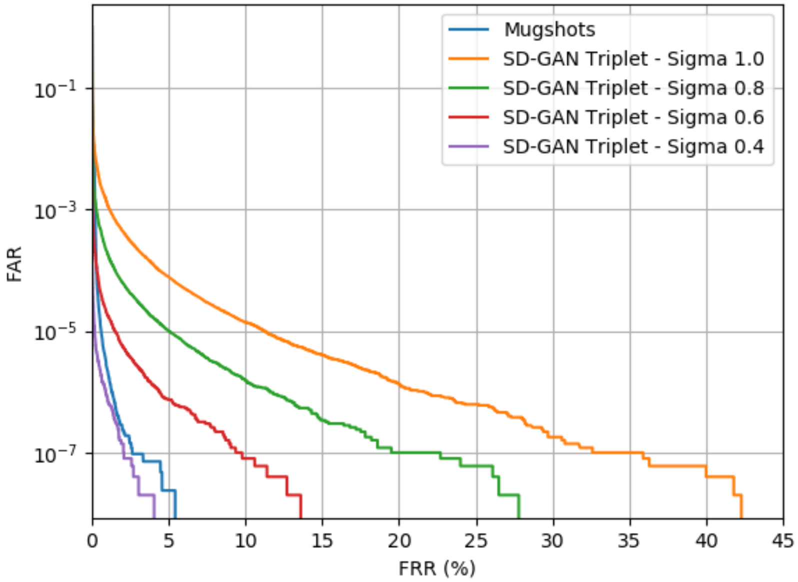
Biometrics Evaluation of GANs

SD-GAN + Triplet loss



Z_{IV} is a Gaussian random vector controlling the intra-class variability

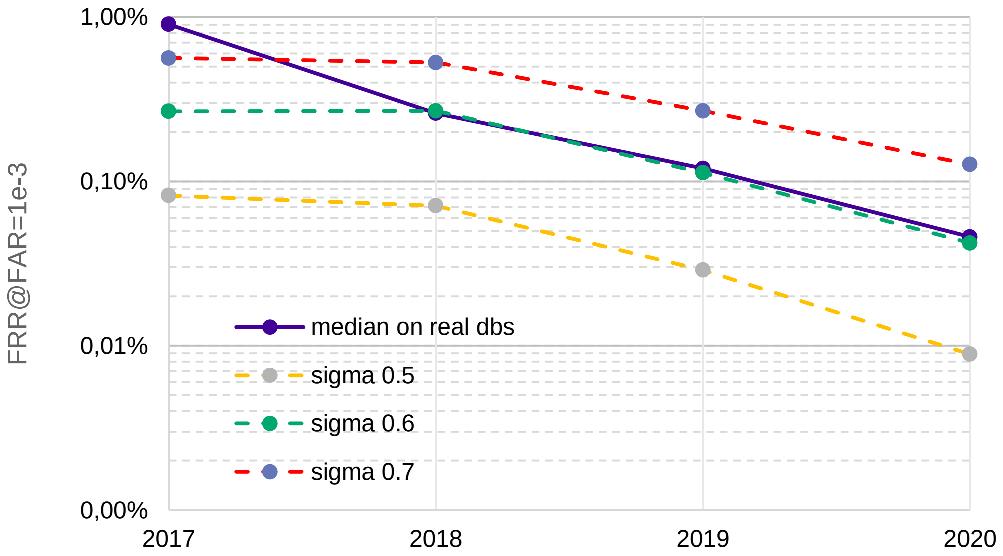
Biometric evaluation with various standard deviations for intra-class distributions



- GANs can be tuned to adjust intra-class variability, in order to reach desired biometric performances.

Evaluation of Biometric Algorithm Improvements

Comparison of biometric performance-evolution on real and synthetic datasets



- Most improvements in biometric algorithms can be seen on synthetic datasets.

Conclusions

1. Overfitting is *not* occurring. New IDs *are* being generated.
2. Synthetic images allow wider system tests than with the replication of a small dataset.
   - Without privacy concerns
   - Non-regression test, speed test, loading test …
   - Moderate control of pose, illumination, age, glasses and gender distributions
3. Synthetic Images can be used to compute some biometric performances.
   - Behaviour with larger gallery
   - Comparison of different algorithms
4. Today, none of the assessed methods was able to fully disentangle identity. It is still a research topic.
   - Not yet ready to be used as training datasets for biometric algorithms

Questions ?

stephane.gentric@idemia.com

References

Deng et.al. (2019)

J. Deng, J. Guo, N. Xue, and S. Zafeiriou. Arcface: Additive angular margin loss for deep face recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pages 4690–4699, 2019.

Donahue et.al. (2018)

C. Donahue, A. Balsubramani, J. McAuley, and Z. C. Lipton. Semantically decomposing the latent spaces of generative adversarial networks. In International Conference on Learning Representations, 2018.

Karras et al. (2018)

T. Karras, T. Aila, S. Laine, and J. Lehtinen. Progressive growing of GANs for improved quality, stability, and variation. In International Conference on Learning Representations, 2018.

Marriott et.al. (2020)

R. Marriott, S. Romdhani, and L. Chen. Taking control of intra-class variation in conditional gans under weak supervision. In 2020 15th IEEE International Conference on Automatic Face and Gesture Recognition (FG 2020), pages 283–290, 2020.

Shen et al. (2020)

Y. Shen, J. Gu, X. Tang, and B. Zhou. Interpreting the latent space of gans for semantic face editing. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pages 9243–9252, 2020.

Marriott et.al. (2020)

R. Marriott, S. Madiouni, S. Romdhani, S. Gentric, and L. Chen. An Assessment of GANs for Identity-related Applications. In 2020 IEEE International Joint Conference on Biometrics (IJCB 2020)