脸部转正！GAN能否让侧颜杀手、小猪佩奇真容无处遁形？

原创 bryant8 机器学习与生成对抗网络 1月7日

欢迎点击上方蓝字，关注啦~

相关阅读：

GAN整整6年了！是时候要来捋捋了！

容颜渐失！GAN来预测？

弱水三千，只取你标！AL（主动学习）结合GAN如何？

异常检测，GAN如何gan ？

虚拟换衣！速览这几篇最新论文咋做的！

脸部妆容迁移！速览几篇用GAN来做的论文

【1】GAN在医学图像上的生成，今如何？

01-GAN公式简明原理之铁甲小宝篇

这日，你漫步于街头，闲看车水马龙、行人熙攘。

忽然，一个惊世骇俗的侧颜闪现于你左前方15米处，你不禁心头一惊：

“这般侧颜，必属盛世无双，人间难得几回观？”

于是你夺步寻去，意欲一览此般倾世容颜……

几分钟后，你大失所望，怨道：

“正脸居然不过如此！真是侧颜杀手！”

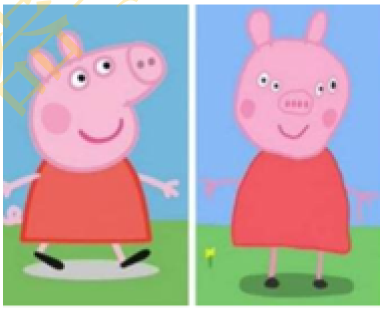
图自网络，侵删。

失落的你、此刻只能打开小猪佩奇动画，一解忧愁。

“这只猪总是用侧脸面对观众，另一面却从来没人见过。奇怪的是，即使从侧颜看去，还是两只眼睛和两个鼻孔。”

于是你抚心、而发出灵魂一问：

“给定侧颜，如何得到其正脸之真容？？？”



哈哈，今天整理的是用GAN进行“**脸部转正**”的论文（待看）。即给定一幅侧脸图像，如何得到正脸图像？

1 (2020-1-6) PI-GAN: LEARNING POSE INDEPENDENT REPRESENTATIONS FOR MULTIPLE PROFILE FACE SYNTHESIS

https://arxiv.xilesou.top/pdf/2001.00645.pdf

我们往往希望，通过某种方法去提取某一人脸姿态图像的固定不变的“姿态表示特征”，并以此去生成多种视图的姿态图像，例如人脸转正（给定一幅侧脸，即可推知正脸图像）。从某个面部朝向姿势图像去合成其它面部姿势，如何获取“姿态表示特征”仍然是一个难题。人脸转正在诸如多媒体安全性、计算机视觉、机器人技术等各个领域都有应用价值。

为了解决这个问题，本文提出了PIGAN（循环共享编码器/解码器框架），利用具有编码器-解码器结构的生成对抗网络（GAN），联合判别器网络，去学习提取“与姿态无关的特征”，再利用其实现逼真的人脸合成。

与传统的GAN相比，它还由辅助的编码器-解码器框架组成，它与主框架中共享权重，并从原始姿势图像重构图像；主框架着重于创建“解耦表示”，而次框架旨在还原原始面孔。使用CFP高分辨率数据集进行方法的验证。

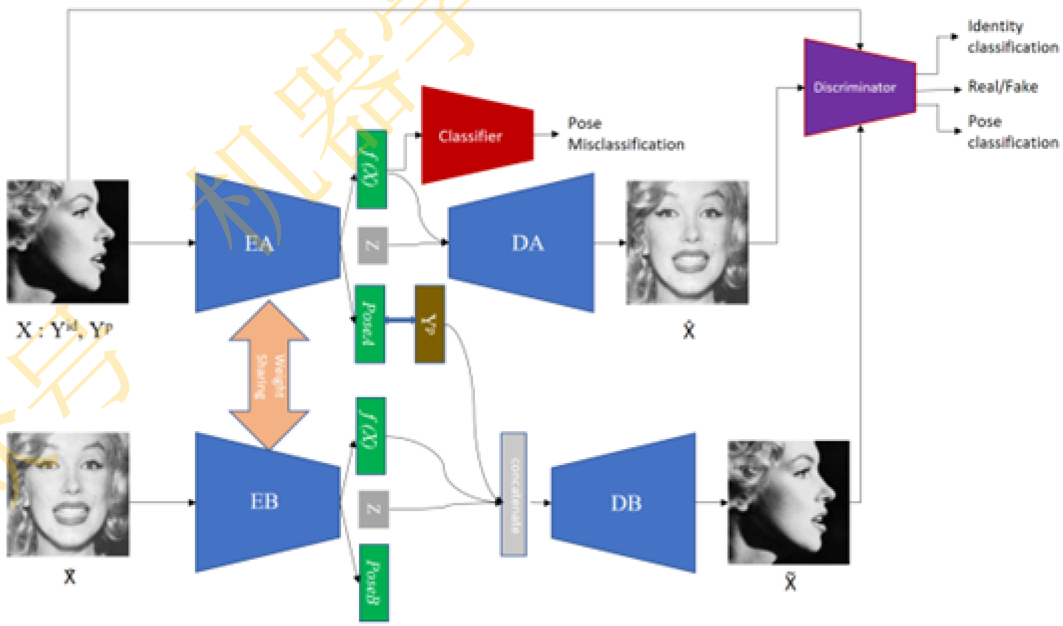


Figure 1: Proposed framework.

实验：CFP数据集。由500个人组成，每人包含10幅正面图像和4幅侧脸。在450个人上训练模型，并对剩余进行评估。基本框架架构遵循DCGAN 的实现。效果图如下，有些乍一看还挺好？

Figure 2: Generated face images. Each pair shows left face as input and right face as generated frontal image.

2 (2019-02-26) BoostGAN for Occlusive Profile Face Frontalization and Recognition

https://arxiv.xilesou.top/pdf/1902.09782.pdf

有许多因素会影响人脸识别效果，例如姿势、遮挡、照明、年龄等等，其中最主要的是大姿势和遮挡问题，这些问题甚至可能模型的性能下降10%以上。

“姿势不变特征表示”和利用生成对抗网络（GAN）进行人脸转正已被广泛用于解决姿势问题。然而，受遮挡的侧脸的识别仍然是一个待解决的问题。为此本文提供一种有效的解决方案，即使面对脸部关键点区域（例如眼睛，鼻子等）受损或遮挡的侧脸图像，也去尝试识别。

具体来说，提出一个BoostGAN，用于去遮挡，正面化和面部识别。基于面部遮挡是部分且不完整的假设，多个遮挡块的图像将作为输入，也就是所谓的的“knowledge boosting”，例如身份和纹理信息。然后，进一步设计了一种新的聚合网络模块，用于最终精细的图像合成。

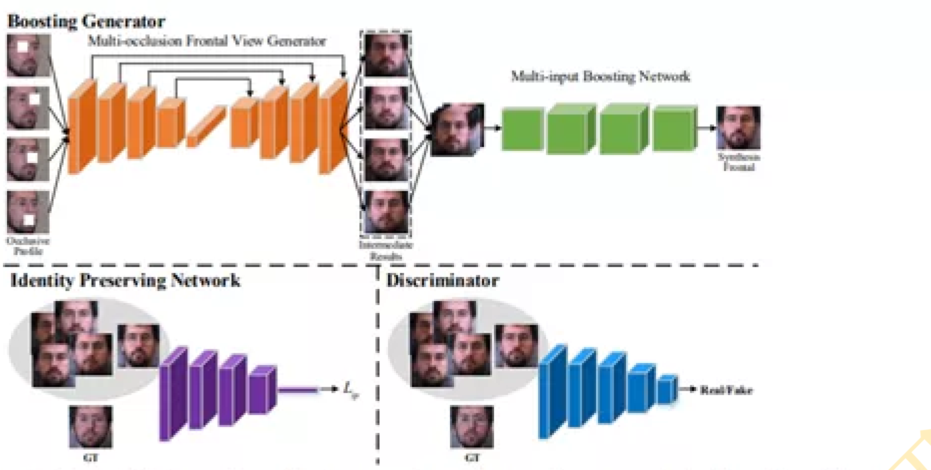
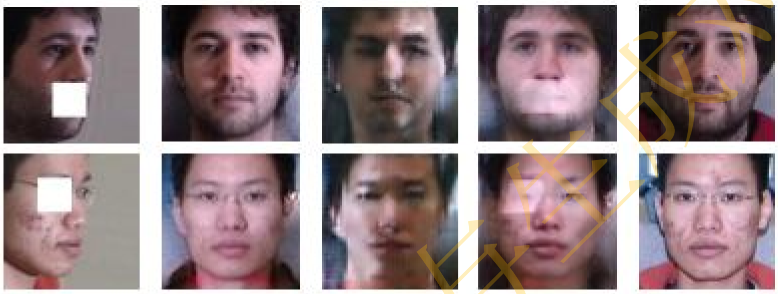


Figure 2. The framework of BoostGAN, in an end-to-end and coarse-to-fine architecture. Two parts are included: multi-occlusion frontal view generator and multi-input boosting network. For the former, the coarse de-occlusion and slight identity preserving images are generated. For the latter, the photo-realistic, clean, and frontal faces are achieved by ensemble complementary information from multiple inputs in a boosting way.



(a) Profile (b) **Ours** (c) [21] (d) [12] (e) GT

Figure 1. Synthesis results by testing the existing models on occluded faces. The poses of the 1^{st} row and the 2^{nd} row are 45° and 60°, respectively. GT denotes the ground truth frontal images.

3 (2018-10-6) Learning a High Fidelity Pose Invariant Model for High-resolution Face Frontalization

https://arxiv.xilesou.top/pdf/1806.08472.pdf

人脸正面化是指根据给定的侧脸去合成人脸正面视图的过程。由于遮挡和扭变，要恢复较好的结果、以高分辨率保存纹理细节极为困难。本文提出了一种高保真姿势不变模型（HF-PIM）来产生逼真的、能保持身份特征一致的结果。

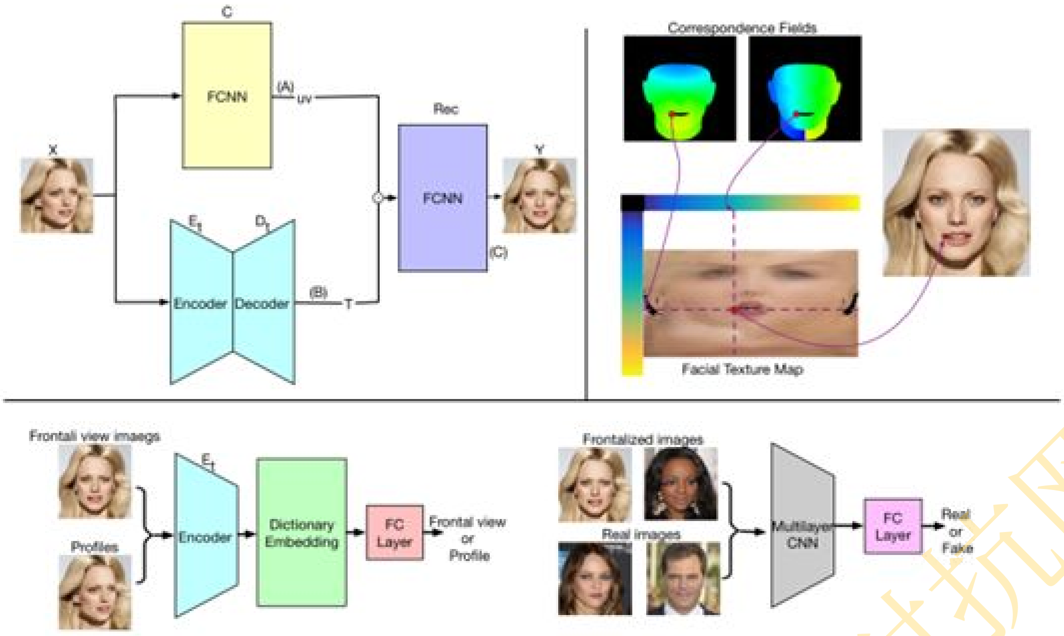


Figure 1: Left side on top: the framework of our HF-PIM to frontalize face images. The procedure consists of correspondence field estimation (A), facial texture feature map recovering (B) and frontal view warping (C). The right side on top: an illustration about the warping procedure discussed in Eq. 1. Those red dots and purple lines indicate the relationships between the facial texture map, correspondence field, and the RGB color image. Bottom side: the discriminators employed for ARDL (on the left) and ordinary adversarial learning (on the right).

Figure 3: High-resolution frontalized results on the testing set of CelebA-HQ. The first row is the input profile images. The second row is the frontalized images produce by our HF-PIM. The results of CAPG-GAN (on the left for each subject) and TP-GAN (on the right) are shown in the third row.

4 (2018-3-4) Load Balanced GANs for Multi-view Face Image Synthesis

https://arxiv.xilesou.top/pdf/1802.07447.pdf

从单个图像去合成多视图人脸是一个ill-posed病态、不定的问题，结果往往有严重的外观失真。生成逼真的、保留身份的多视图仍然是一个挑战。本文提出了负载平衡生成对抗网络（LB-GAN），可以将输入人脸图像的偏航角精确地旋转到任意指定角度。

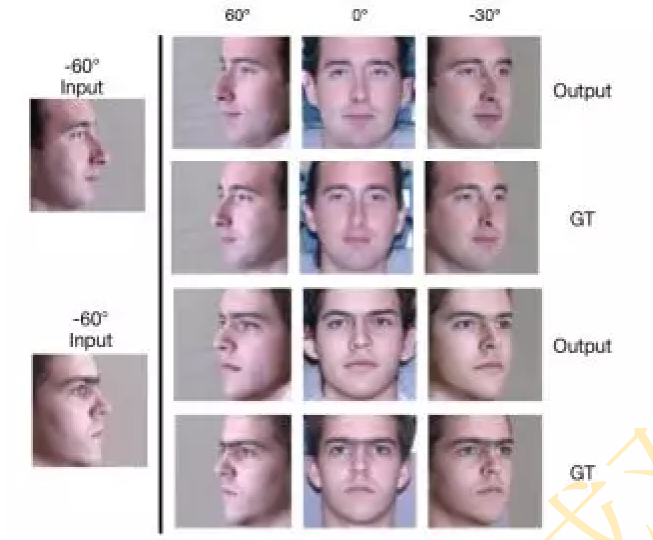


Figure 1: Face rotation results by our LB-GAN. According to the degrees on the top for each column on the right-hand, the inputs are rotated to a specified pose. GT stands for the ground truth.

LBGAN将具有挑战性的综合问题分解为两个子任务：人脸标准化和人脸编辑。标准化首先对输入图像进行正面化，然后编辑器将正面化的图像旋转到所需姿势。为了生成逼真的局部细节，对标准化器和编辑器进行两阶段训练，并通过有条件的self-cycle loss和基于L2 loss的attention进行约束。

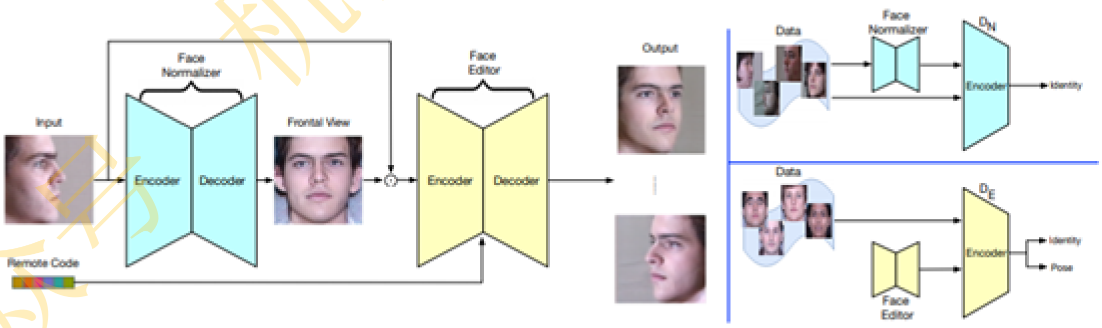


Figure 2: The framework of our LB-GAN for multi-view face image synthesis.

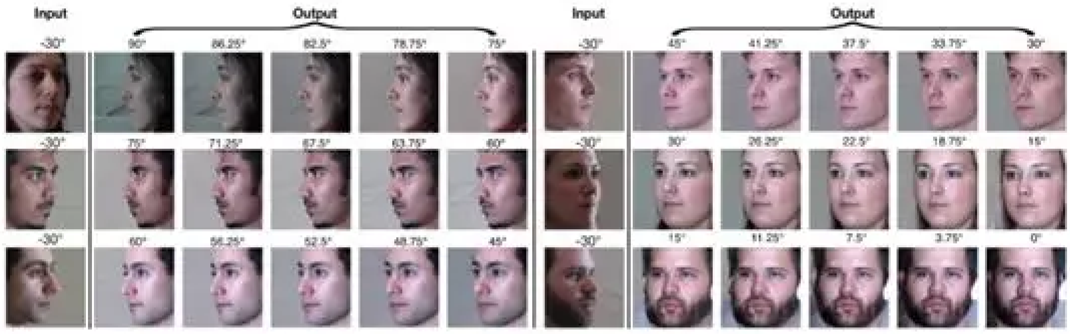


Figure 3: Face rotation results on Multi-PIE. Each input is rotated to the specified yaw angle which is indicated by the degree above the output.

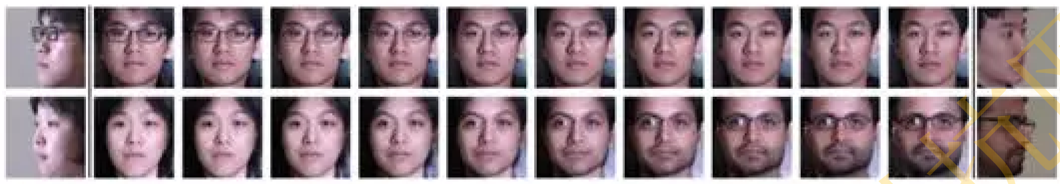


Figure 4: Synthesizing new samples through interpolating identity representations. For each row, the samples in the middle are synthesized by the interpolation of the identity representations of the far left and the far right faces.

5 (2017-12-13) UV-GAN Adversarial Facial UV Map Completion for Pose-invariant Face Recognition

https://arxiv.xilesou.top/pdf/1712.04695.pdf

最近提出的鲁棒的3D人脸对齐方法在3D人脸模型和2D人脸图像之间建立了密集或稀疏的对应关系。这些方法的使用对于面部纹理分析既是挑战也是机遇。特别是，通过使用拟合模型（fitted model）对图像进行采样，可以创建面部UV。D但由于遮挡，UV图总是不完整的。本文提出UV-GAN，用生成的UV图生成任意姿态的2D面部图像。

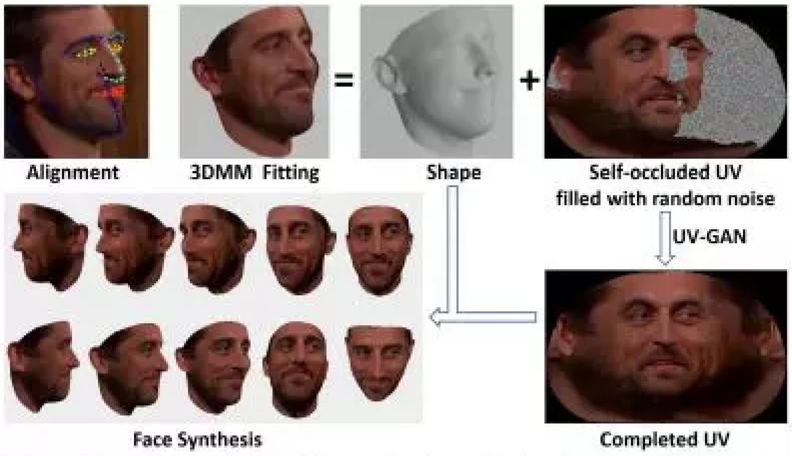


Figure 1. Adversarial UV completion. After fitting a 3DMM to the image, we retrieve a 3D face with an incomplete UV map. We learn a generative model to recover the self-occluded regions. By rotating a 3D shape with complete UV map, we can generate 2D faces of arbitrary poses, which can either augment pose variations during training or narrow pose discrepancy during testing for pose-invariant face recognition.

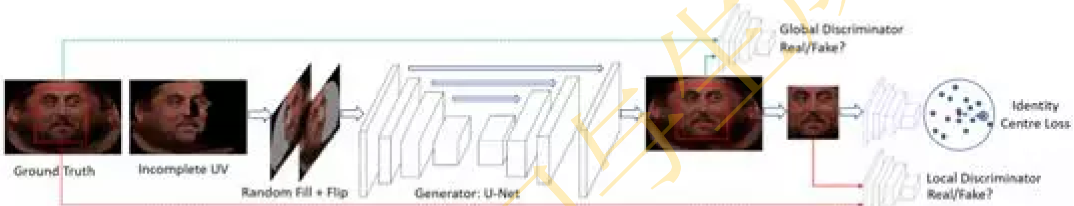


Figure 2. Network architecture. It consists of one generator and two discriminators. The generator takes the incomplete UV map as input and outputs the full UV map. Two discriminators are learned to validate the genuineness of the synthetic UV texture and the main face region. The pre-trained identity classification network that would always be fixed, is to further ensure the newly generated faces preserve the identity. Note that only the generator is required for the testing stage.

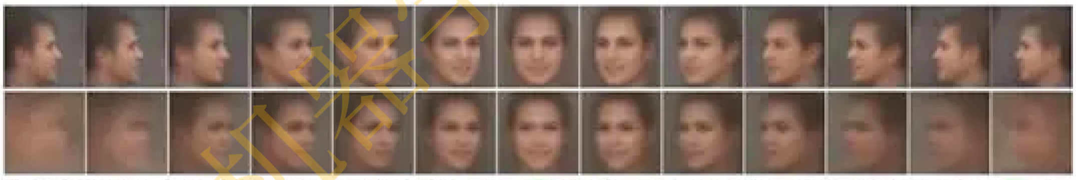


Figure 3. The mean faces of 13 pose groups in the CASIA dataset [38]. The first row is our results, and the second row is from [33]. Our mean faces are more clear under large pose variations, and there is only slight blurriness on facial organs even for profile faces.

6 (2017-08-17) Towards Large-Pose Face Frontalization in the Wild

https://arxiv.xilesou.top/pdf/1704.06244.pdf

尽管最近在使用深度学习的面部识别方面取得了进步，但是在较大的姿势变化下，严重影响性能。学习姿势不变特征是一种解决方案，但是需要昂贵的大规模数据标注和精心设计的特征学习算法。本文结合 3D Morphable Model (3DMM) 和 GAN 来进行人脸转正，称为FF-GAN。

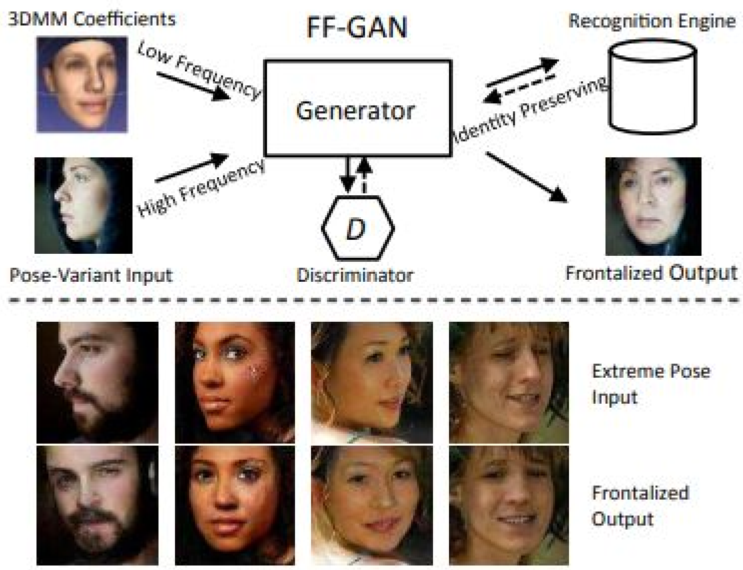


Figure 1: The proposed FF-GAN framework. Given a non-frontal face image as input, the generator produces a high-quality frontal face. Learned 3DMM coefficients provide global pose and low frequency information, while the input image injects high frequency local information. A discriminator distinguishes generated faces against real ones, where high-quality frontal faces are considered real. A face recognition engine is used to preserve identity information. The output is a high quality frontal face that retains identity.

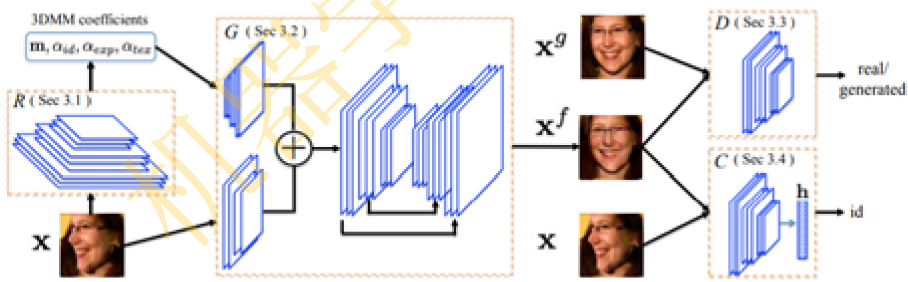


Figure 2: The proposed FF-GAN for large-pose face frontalization. R is the reconstruction module for 3DMM coefficients estimation. G is the generation module to synthesize a frontal face. D is the discrimination module to make real or generated decision. C is the recognition module for identity classification.

7 (2017-08-3) Beyond Face Rotation: Global and Local Perception GAN for Photorealistic and Identity Preserving Frontal View Synthesis

https://arxiv.xilesou.top/pdf/1704.04086.pdf

本文提出了一种两路生成对抗性网络（TP-GAN），用于通过同时感知全局结构和局部细节来实现逼真的正面视图合成。除了常用的全局编码器/解码器网络之外，还提出了四个局部块网络来处理局部纹理。此外，引入对抗性损失，对称性损失和身份保留损失的损失组合。

Figure 1. Frontal view synthesis by TP-GAN. The upper half shows the 90° profile image (middle) and its corresponding synthesized and ground truth frontal face. We invite the readers to guess which side is our synthesis results (please refer to Sec. 1 for the answer). The lower half shows the synthesized frontal view faces from profiles of 90°, 75° and 45° respectively.

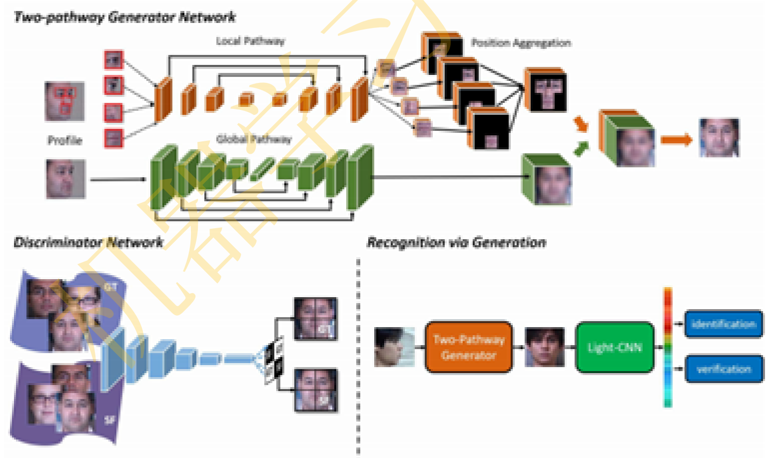
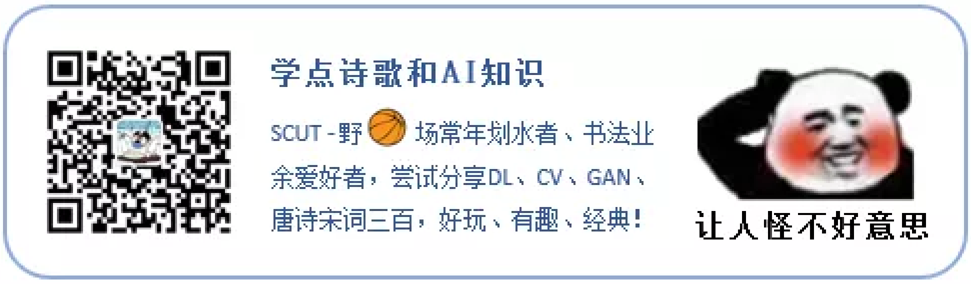


Figure 2. General framework of TP-GAN. The Generator contains two pathways with each processing global or local transformations. The Discriminator distinguishes between synthesized frontal (SF) views and ground-truth (GT) frontal views. Detailed network architectures can be found in the supplementary material.

更多分享欢迎关注本公众号：


学点诗歌和AI知识
SCUT - 野 场常年划水者、书法业余爱好者，尝试分享DL、CV、GAN、唐诗宋词三百，好玩、有趣、经典！
让人怪不好意思