有点夸张、有点扭曲！速览这些GAN如何夸张漫画化人脸！

原创 bryant8 机器学习与生成对抗网络 1月21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！

天降斯雨，于我却无！GAN用于去雨如何？

脸部转正！GAN能否让侧颜杀手、小猪佩奇真容无处遁形？

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

这次整理的是，用GAN将人脸夸张漫画化的方向！

1 (2018-07-24) Unpaired Photo-to-Caricature Translation on Faces in the Wild

https://arxiv.xilesou.top/pdf/1711.10735.pdf

Unpaired Photo-to-Caricature Translation on Faces in the Wild

Ziqiang Zheng[a], Chao Wang[a], Zhibin Yu[a], Nan Wang[a], Haiyong Zheng[a,*], Bing Zheng[a]

[a]No. 238 Songling Road, Department of Electronic Engineering, Ocean University of China, Qingdao, China

条件生成对抗网络（cGAN）使得图像间的转换取得了很大的进步。一些基于循环一致性损失的、无需配对训练数集的方法，例如DualGAN，CycleGAN和DiscoGAN确实很受欢迎。但是，对于需要高级视觉信息转换的翻译任务来说，例如从普通照片到极具讽刺、夸张和艺术形变性的漫画的转换仍然是非常具有挑战性的。

本文提出了一种基于学习的方法来解决此类问题。为了在转换时，兼顾局部统计量和全局结构，设计了一个带有粗区分和一个细区分判别器的双路模型。对于生成器，使用了感知损失，对抗损失和一致性损失，以实现两个不同领域的表示学习。另外，可以通过辅助噪音输入来了解风格。

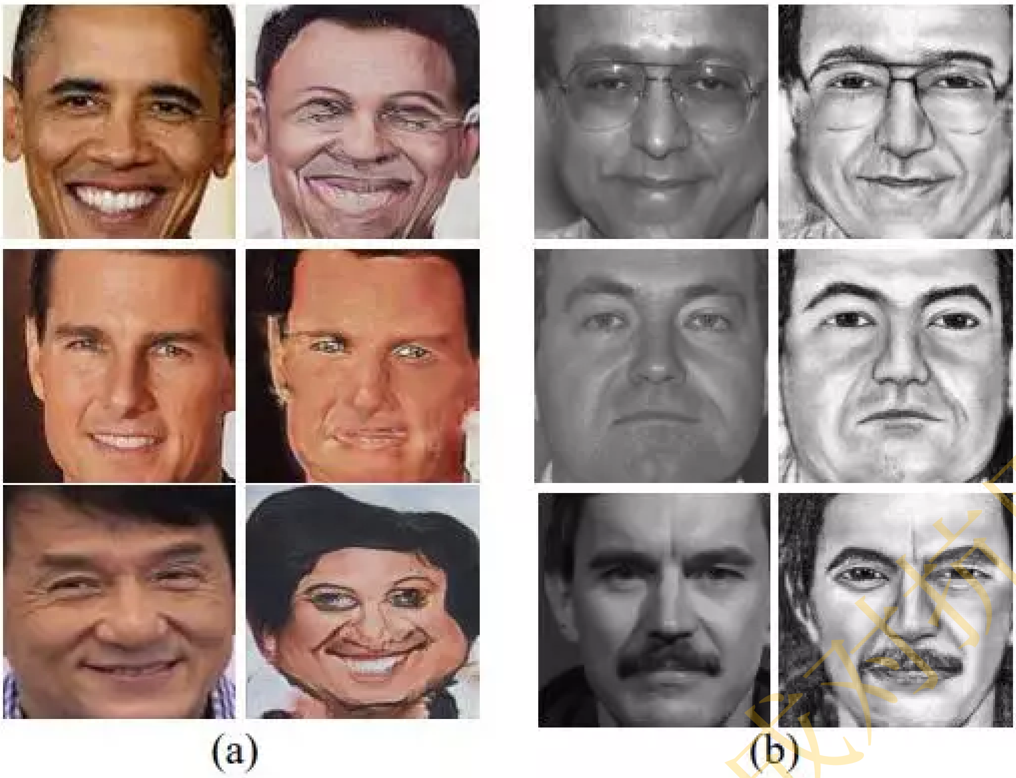


Figure 1: Translating faces in the wild from photo to caricature with different styles by our proposed method. (a) Example results on IIIT-CFW dataset [25]; (b) Example results on PHOTO-SKETCH dataset [26, 27].

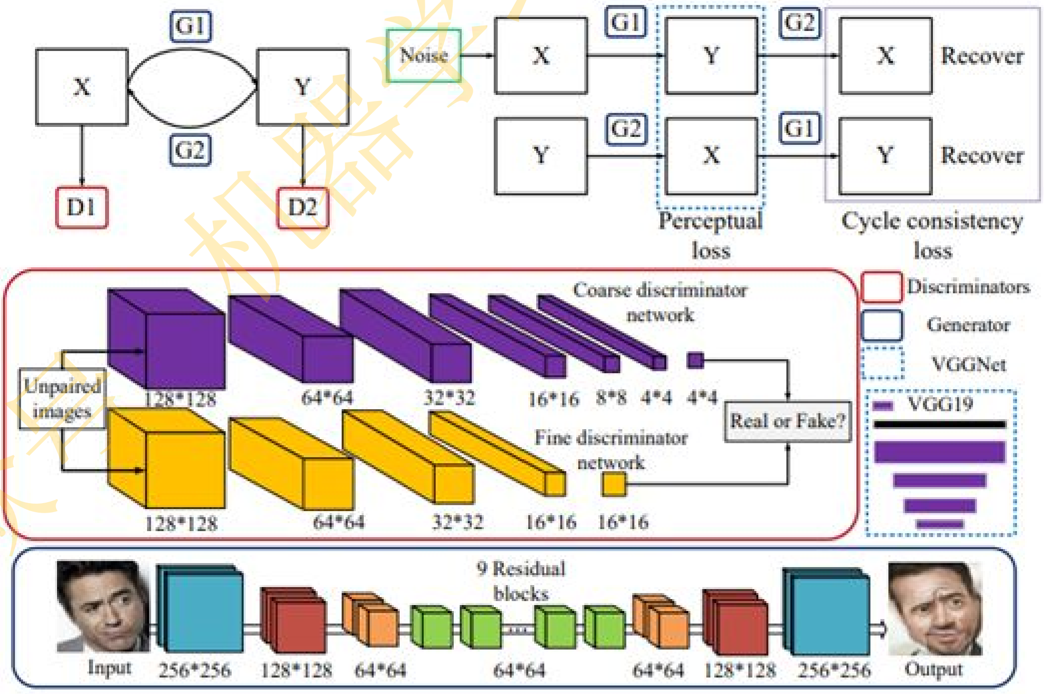


Figure 2: Network architecture and data flow chart of our proposed method for face photo-to-caricature translation.

Figure 4: Comparison of state-of-the-art image-to-image translation methods with our proposed method for face photo-to-caricature translation on IIIT-CFW-P2C dataset.

2 (2018-11-1) CariGANs Unpaired Photo-to-Caricature Translation

https://arxiv.xilesou.top/pdf/1811.00222.pdf

CariGANs: Unpaired Photo-to-Caricature Translation

KAIDI CAO*†, Tsinghua University
JING LIAO‡, City University of Hong Kong, Microsoft Research
LU YUAN, Microsoft AI Perception and Mixed Reality

(a) Photo (b) Hand-drawn (c) Gatys et al. [2015] (d) Zhu et al. [2017a] (e) Ours (with ref) (f) Ours (with noise)

Fig. 1. Comparison of the caricature drawn manually (b) and generated automatically with neural style transfer [Gatys et al. 2015] (c), CycleGan [Zhu et al. 2017a] (d), and Our CariGANs with a given reference (e) or a random noise (f). Please note networks used in (d)(e)(f) are trained with the same dataset. And the reference used in the result is overlaid on its bottom-right corner. Photos: MS-Celeb-1M dataset, hand-drawn caricatures (from top to bottom): ©Lucy Feng/deviantart, ©Tonio/toonpool.

人脸漫画化是一种传达夸张、幽默或讽刺意味的艺术表示形式。本文提出了一个无需成对训练数据集的照片到漫画转换的生成对抗网络（GAN）：CariGAN。

使用两个模块来显式地建模几何夸大和外观风格化：一个是CariGeoGAN，仅对几何形变上进行建模，即从面部照片到漫画的几何转换；另一个是CariStyGAN，在风格的外观层面上将漫画风格转移到面部照片。这样，一个困难的跨域转换问题被分解为两个更简单的任务。

与先进的方法相比，CariGAN生成的漫画更接近于手绘，同时更好地保持原有人脸的个性特征。此外，还允许用户控制形状的夸张程度和变化，或给出示例漫画来生成相应的风格。

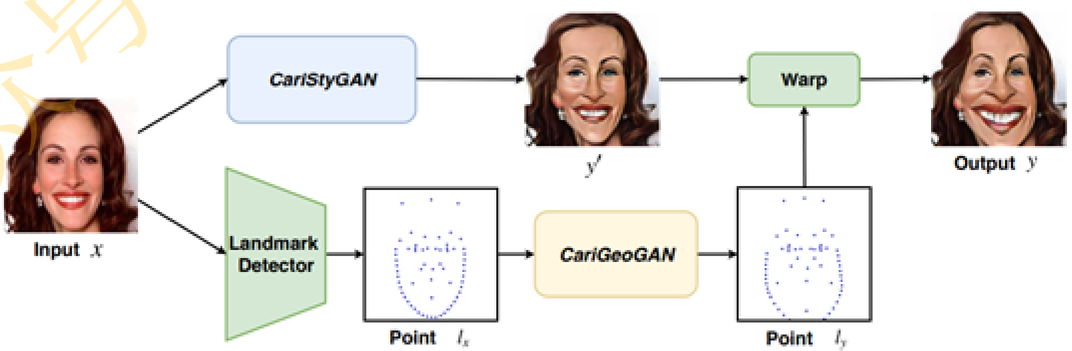


Fig. 2. Overall Pipeline of Proposed Method. Input image: CelebA dataset.

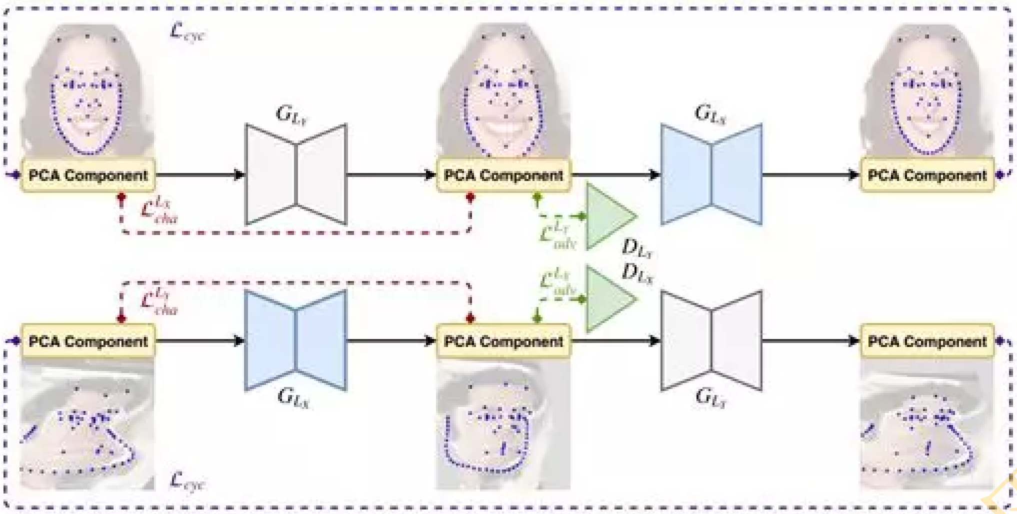


Fig. 4. Architecture of CariGeoGAN. It basically follows the network structure of CycleGAN with cycle Loss $\mathcal{L}_{cyc}$ and adversarial loss $\mathcal{L}_{gan}$. But our input and output are vectors instead of images, and we add a characteristic loss $\mathcal{L}_{cha}$ to exaggerate the subject's distinct features. Input images: CelebA dataset.

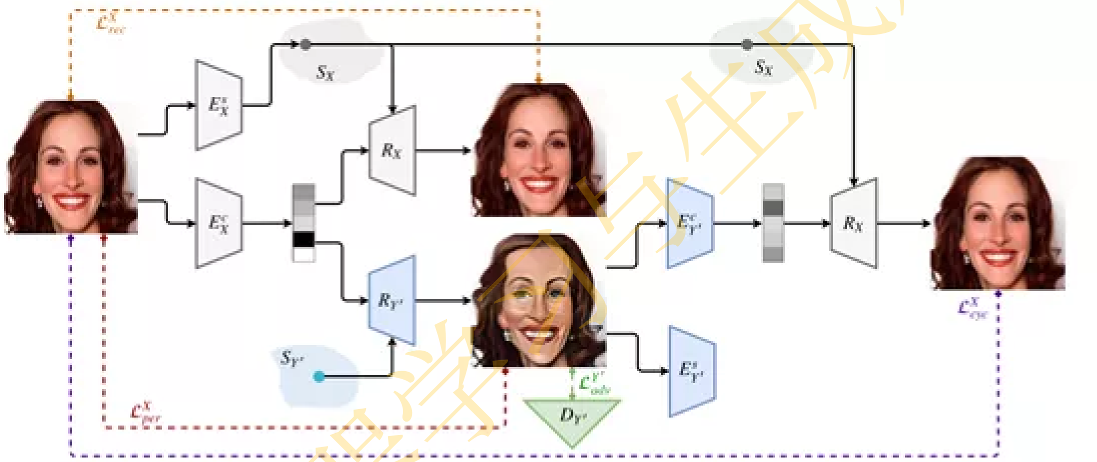


Fig. 8. Architecture of our CariStyGan. For simplicity, here we only show the network architecture for the translation $X \to Y'$. And the network architecture for the reverse translation $Y' \to X$ is symmetric. Input image: CelebA dataset.

3 (2018-11-20) CariGAN Caricature Generation through Weakly Paired Adversarial Learning

https://arxiv.xilesou.top/pdf/1811.00445.pdf

CariGAN: Caricature Generation through Weakly Paired Adversarial Learning

Wenbin Li[†*], Wei Xiong[‡*], Haofu Liao[‡], Jing Huo[†], Yang Gao[†], Jiebo Luo[‡]
[†] State Key Laboratory for Novel Software Technology, Nanjing University, China
[‡] Department of Computer Science, University of Rochester, USA
liwenbin.nju@gmail.com, {wei.xiong, haofu.liao}@rochester.edu
{huojing, gaoy}@nju.edu.cn, jluo@cs.rochester.edu

传统的漫画生成方法主要使用low-level的几何变换（例如图像扭曲）来生成夸张的漫画图像，这些图像在内容和风格方面缺乏丰富性和多样性。尽管生成对抗网络（GAN）使得图像间的转换成为可能，但由于漫画这种图像数据分布的巨大差异，导致基于GAN的模型应用于此任务上的效果不尽人意。

本文将漫画生成建模为一个只需要弱配对训练数据集的图像转换任务，提出CariGAN来解决问题。具体地，为了强制进行合理的夸张和面部变形，采用面部特征点作为附加条件来约束所生成的图像，并设计了一种图像融合机制来鼓励模型将注意力集中在面部关键部位上，以便可以在这些区域中生成更生动的细节。同时，提出了一种多样性损失，以鼓励产生多样化的结果，以帮助减轻常规基于GAN的模型的“模式崩溃”问题。在大规模“ WebCaricature”数据集上的实验表明，与最新模型相比，CariGAN可以生成更多具有更多样性的漫画。

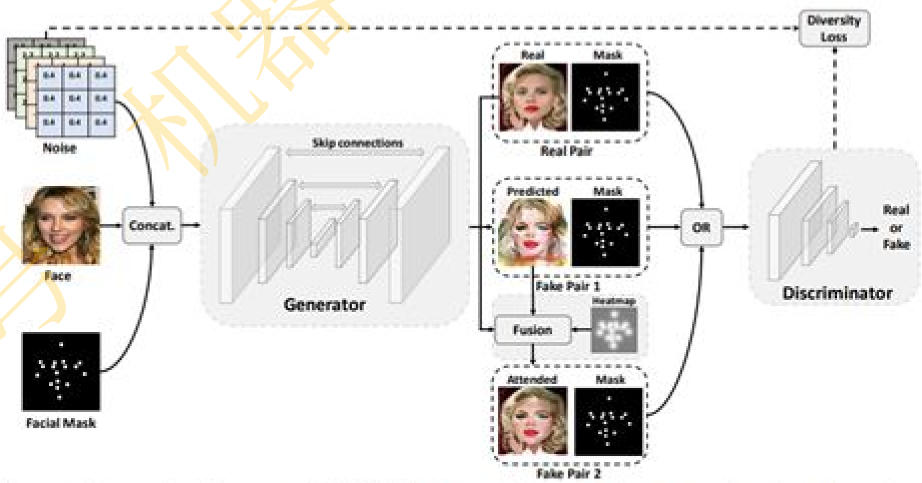


Figure 2. The overall framework of the proposed CariGAN. The input of the generator is a concatenation of a random noise map, a face image and a facial mask. The input data is fed into a U-net generator to generate a fake caricature. Then, using our *image fusion mechanism*, we combine the ground-truth and the generated fake caricature to get an additional fused fake caricature. These caricatures are concatenated with the input facial mask, respectively, and fed into the discriminator. The discriminator then tries to distinguish the real from the fakes. In addition to the adversarial loss, a *diversity loss* is proposed to constrain the outputs to be more diverse in style and content.

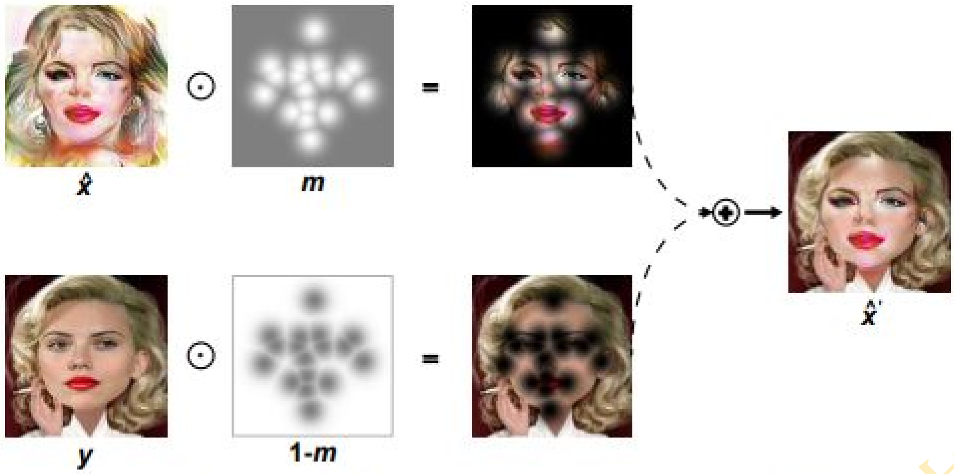


Figure 3. Illustration of the image fusion procedure.

在模型中，随机噪声控制着图像的颜色和样式等。但实际上，提出的模型可能会“模式崩溃”，即输入噪声可能不会影响最终结果。为解决“模式崩溃”问题，提出了一种多样性损失，以迫使模型生成具有更大多样性的图像。基本思想是：假设由两种不同噪声（但具有相同的输入面部和脸部mask）生成的两种漫画图像之间的差异，是这两种噪声之间的差异的线性函数。

例如为生成器提供了一个人脸图像x和一个二进制mask，但具有两个不同的噪声z1和z2。生成器针对这两个输入分别输出两漫画，即xˆ1和xˆ2。有：xˆ1 = G（x，p，z1），xˆ2 = G（x，p，z2）。而鉴别器D的最后一个卷积层中提取这两个漫画的特征。将提取的特征表示为f1 = D（xˆ1，p），f2 = D（xˆ2，p）。

提取的特征实际上暗含生成图像的身份，姿势和样式等信息。但由于这两个特征是从具有相同人脸的两个假漫画中提取的，因此将这两个特征之间的差异视为风格和其他不重要属性之间的差异是合理的。从而将两个特征之间的差异强制为两个输入噪声之间的差异的线性函数。如此一来，生成的的多样性可以通过输入噪声明确控制：

$$\mathcal{L}_{div} = \left(\frac{\|f_1 - f_2\|_2^2}{\|f_1\|_2^2 + \|f_2\|_2^2} - \frac{\|z_1 - z_2\|_2^2}{\|z_1\|_2^2 + \|z_2\|_2^2} \right)^2$$

Figure 5. Interpolating on the noise z. The first and second columns are the input images and ground truth, respectively. The rest are the outputs with linearly interpolated noises. Odd rows: our model without diversity loss. Even rows: our model with diversity loss. Please pay special attention to the colors and textures of the generated caricatures.

4 (2019-04-16) WarpGAN Automatic Caricature Generation

https://arxiv.xilesou.top/pdf/1811.10100.pdf

WarpGAN: Automatic Caricature Generation

Yichun Shi* Debayan Deb* Anil K. Jain
Michigan State University, East Lansing MI 48824
{shiyichu, debdebay}@msu.edu, jain@cse.msu.edu

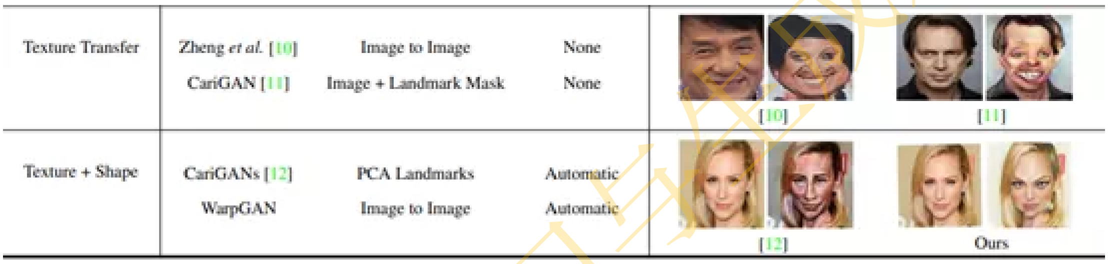
(a) Photo (b) WarpGAN (c) WarpGAN (d) Artist (e) Artist

Figure 1: Example photos and caricatures of two subjects in our dataset. Column (a) shows each identity's real face photo, while two generated caricatures of the same subjects by WarpGAN are shown in column (b) and (c). Caricatures drawn by artists are shown in the column (d) and (e).

本文提出WarpGAN，一种全自动网络，可在输入面部照片的情况下生成漫画。除了能够迁移丰富的纹理风格外，WarpGAN还学会自动预测一组控制点，这些控制点可以将照片扭曲成漫画，同时保留人脸身份信息。

本文引入了一种保持身份信息的对抗性损失，以帮助判别器区分不同的个人。此外，WarpGAN可通过控制夸张程度和视觉风格来定制生成的漫画类型。在公共领域数据集WebCaricature上的实验结果表明，WarpGAN能够生成漫画，这些漫画不仅保留身份，而且为每张输入的照片输出一组多样化的漫画。经过五位漫画专家认为，WarpGAN产生的漫画在视觉上与手绘漫画相当逼近。

与前面几种方法的比较：



Texture Transfer	Zheng *et al.* [10]	Image to Image	None	[10] [11]
	CariGAN [11]	Image + Landmark Mask	None	
Texture + Shape	CariGANs [12]	PCA Landmarks	Automatic	[12] Ours
	WarpGAN	Image to Image	Automatic	

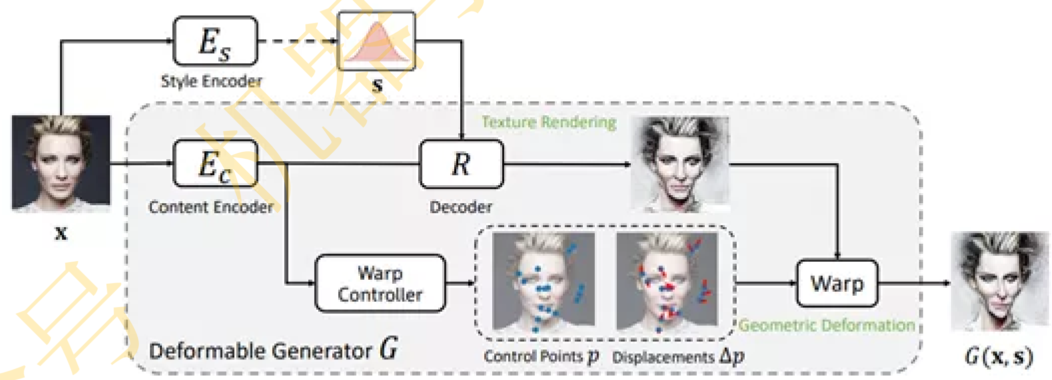


Figure 3: The generator module of WarpGAN. Given a face image, the generator outputs an image with a different texture style and a set of control points along with their displacements. A differentiable module takes the control points and warps the transferred image to generate a caricature.

5 (2020-01-07) MW-GAN: Multi-Warping GAN for Caricature Generation with Multi-Style Geometric Exaggeration

https://arxiv.xilesou.top/pdf/2001.01870.pdf

MW-GAN: Multi-Warping GAN for Caricature Generation with Multi-Style Geometric Exaggeration

Haodi Hou, Jing Huo, Jing Wu, Yu-Kun Lai, and Yang Gao

本文提出Multi-Warping GAN（MW-GAN），包括分别用于进行风格转换和几何夸张形变的网络。通过双向设计去架起图像风格、脸部特征点与相应的潜码空间之间的“桥梁”，生成具有任意风格和几何夸张程度的漫画。此外，将保留身份的损失同时应用于图像空间和脸部特征点空间，从而极大地提高了所生成漫画的质量。实验表明，与现有方法相比，MW-GAN生成的漫画具有更好的质量。

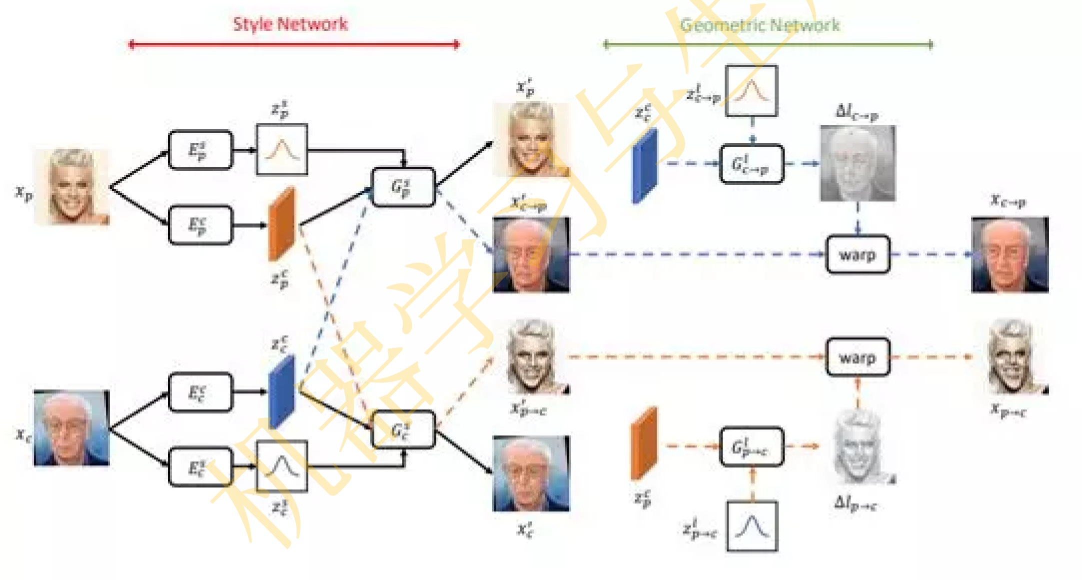


Fig. 1. The network architecture of the proposed multi-warping GAN. The left part is the style network and the right part is the geometric network. The black solid arrows denote the flow of two auto-encoders, with the upper one being the auto-encoder of photos and the lower one for caricature reconstruction. The orange dashed lines denote the flow of photo to caricature generation and the blue dashed lines for caricature to photo generation. In our dual way framework, we assume the caricature and the photo share the same content feature space but have separate style spaces. E_p^c and E_c^c are two encoders to encode the content of photos and caricatures respectively. Similarly defined, E_p^s and E_c^s are two encoders to encode the style of photos and caricatures. Gaussian distribution is imposed on their outputs z_p^s and z_c^s, so that we can sample style codes from Gaussian when translating photos to caricatures or vice versa. To model the geometric transformation procedure, we assume the geometric transformation can be learned by a network ($G_{p\to c}^l$ or $G_{c\to p}^l$) with the image's content code (z_p^c or z_c^c) as input. Besides, by adding a random code ($z_{p\to c}^l$ or $z_{c\to p}^l$) which represents different geometric transformation style, the network can output various geometric transformations for the input image. With the learned geometric transformation, by warping the image with its texture style changed, we get the generated caricature or photo.

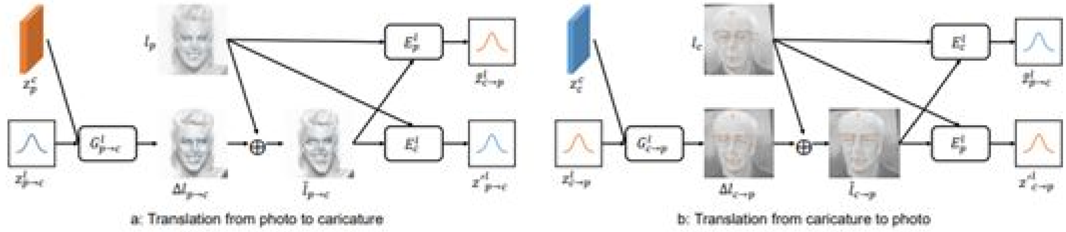


Fig. 2. Geometric Network. The left part is the network for learning a transformation from a photo's landmarks to a caricature's landmarks. The right part is the network for the reverse transformation. For the left network, a generator $G^l_{p\to c}$ with the content code of a photo z^c_p and a landmark transformation latent code $z^l_{p\to c}$ (which can be randomly sampled from a Gaussian distribution) as input will output landmark displacement vectors $\Delta l_{p\to c}$. By adding the displacement vectors to the photo's landmark positions, we get the transformed caricature landmarks $\tilde{l}_{p\to c}$. To make the randomly sampled $z^l_{p\to c}$ correlate to meaningful shape transformation styles, we introduce two encoders and force cycle consistency loss on the encoded latent code and sampled latent code. For example, $z'^l_{p\to c} = E^l_c(\tilde{l}_{p\to c}, l_p)$ is the encoded latent code, we will force $z'^l_{p\to c}$ to be as close as possible to $z^l_{p\to c}$.

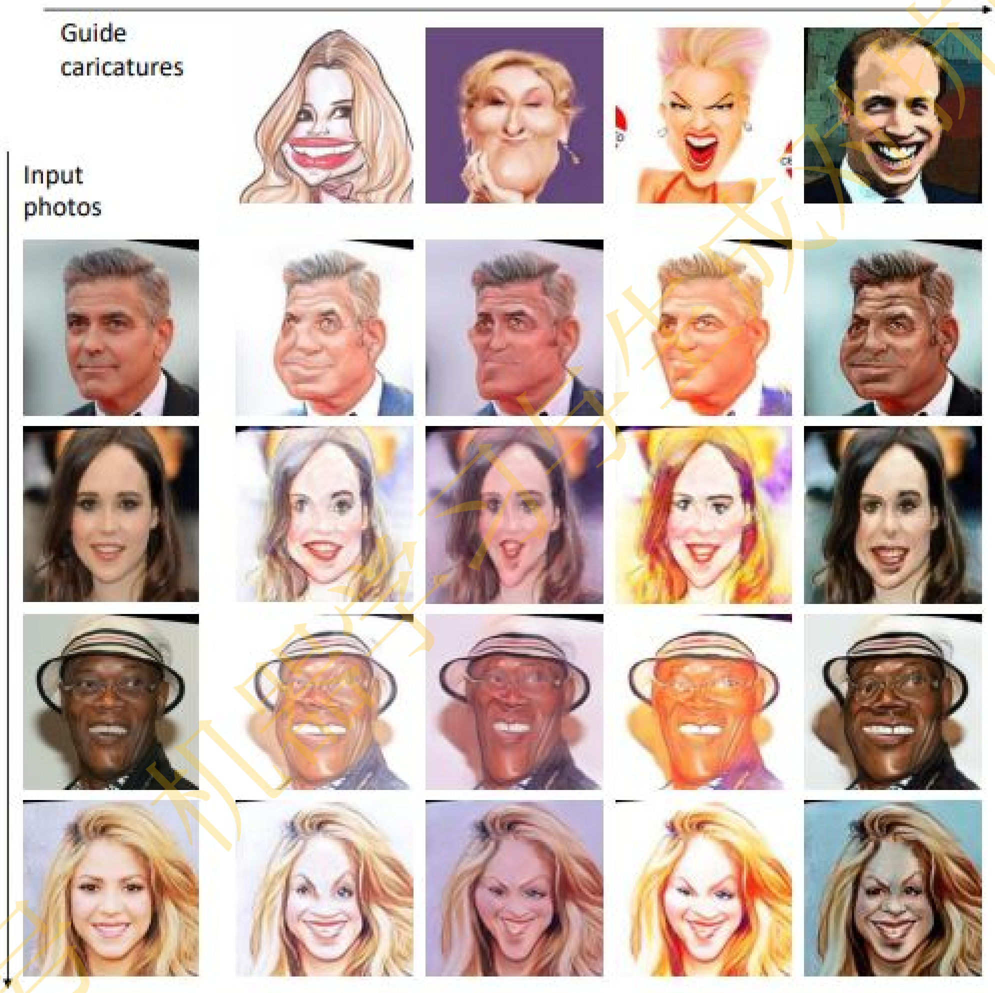
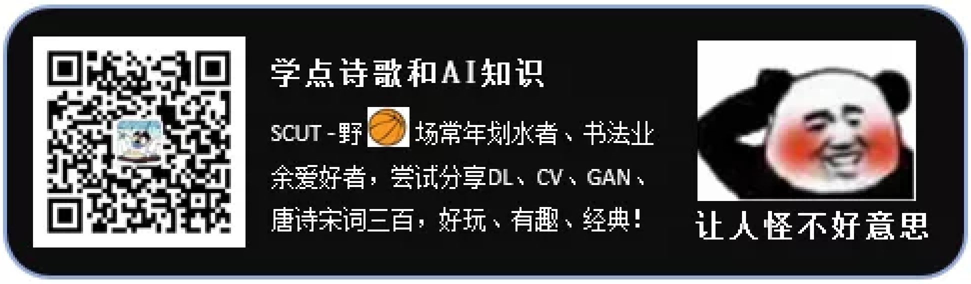


Fig. 7. Sample-guided caricature generation of the proposed method.

先到这吧~（**整理不易，劳烦记得点个在看啦~**

也欢迎关注本公众号~


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