

# 基于对抗生成网络的人脸照片去网纹技术\*

张柯<sup>1,3</sup>, 白富生<sup>1,3</sup>, 吴至友<sup>1,3</sup>, 皮家甜<sup>2,3</sup>, 赵立军<sup>3,4</sup>

(1. 重庆师范大学 数学科学学院; 2. 重庆师范大学 计算机与信息科学学院, 重庆 401331;

3. 智慧金融与大数据分析重庆市重点实验室, 重庆 401331; 4. 马上消费金融股份有限公司, 重庆 401331)

**摘要:**【目的】研究修复带网纹遮挡的人脸照片,以有助于提高后续的人脸验证的准确性。【方法】基于深度学习的模型(针对修复网纹遮挡的人脸照片)在训练时都要求输入对应的网纹数据,但是在实际应用中要获得对应的网纹数据却是非常困难的。为了避免使用对应的网纹数据对人脸图像进行有效的修复,提出了分离对抗生成网络。该网络利用网纹遮挡的人脸照片与干净的人脸照片做像素差生成残差网纹,利用1个分离网络和3个判别网络进行图像修复。【结果】实验结果表明所提出的方法对于消除人脸图像中网纹遮挡有效。【结论】针对带网纹结构遮挡的人脸图像,在对应网纹数据缺失的情况下,通过分离对抗生成网络,依然可以取得很好的图像修复效果。

**关键词:**生成对抗网络;人脸图像修复;去除网纹

**中图分类号:**TP181

**文献标志码:**A

**文章编号:**1672-6693(2019)06-0110-09

人脸图像修复技术<sup>[1]</sup>受到深度学习<sup>[2-3]</sup>特别是卷积神经网络 ConvNet<sup>[4]</sup>技术的影响在近些年取得了非常重大的进展。自从文献<sup>[5]</sup>的开创性工作以来,ConvNets 不断刷新 LFW 性能测评<sup>[6]</sup>的新记录,甚至超越了人类的表现<sup>[7]</sup>。这些新的进展促进了一些具有特殊任务的模型的发展,比如身份证照片与生活照之间做人脸验证<sup>[8]</sup>。在这项特殊的任务中,需要在身份证照片与生活照之间做人脸验证(Face verification between ID photos and life photos, FVBIL)。FVBIL 引起了很多关注,因为它在实际生活中具有非常广泛的应用,如火车站、机场安检、远程开立银行账户等。

目前中国的身份证照片是由中国公安机关提供的。公安机关在提供这些照片时会使用随机的网纹数据给身份证照片进行遮挡以防止身份证在别的地方滥用,如图1所示。在贴上这些随机的网纹后,身份证照片与生活照之间的验证准确率大大降低,为人脸验证带来了巨大的挑战。因此,把损坏了的图像进行修复对人脸验证起着至关重要的作用。但是从被网纹遮挡的图像中恢复出干净清晰的图像是一件非常困难的事情,一方面完全不知道图像中受到遮挡的部分下面原本的图像信息,只能做盲修复,另一方面遮挡图像的网纹是随机产生的,难以获得网纹的生成方式。因此,这在计算机视觉领域中仍然是一个颇具挑战的问题。

对于人脸图像修复而言目前有许多方法,其中基于深度学习的方法在近些年取得了可喜的成果<sup>[1,9-12]</sup>,但是大部分都没有考虑像网纹这样带结构的遮挡。DF-GAN<sup>[13]</sup>和 DeMeshNet<sup>[14]</sup>考虑了网纹这样遮挡的结构信息,其中 DF-GAN 在 Cycle-GAN<sup>[15]</sup>的网络结构基础上多加了1个域,

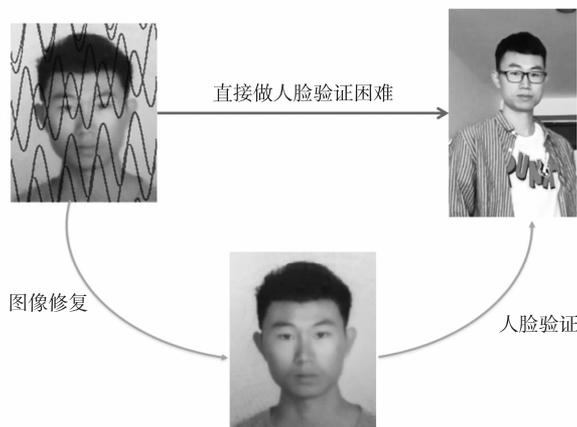


图1 网纹遮挡的人脸图以及人脸验证流程  
Fig. 1 Mesh face and face verification process

\* 收稿日期:2019-06-28 修回日期:2019-07-25 网络出版时间:2019-11-25 10:35

资助项目:重庆市自然科学基金面上项目(No. cstc2019jcyj-msxmX0368);重庆师范大学横向课题“用于人脸比对的身份证网纹照片修复技术”

第一作者简介:张柯,男,研究方向为深度学习与最优化理论与算法,E-mail:15700351784@163.com;通信作者:白富生,男,教授,博士,E-mail:fsbai@cqnu.edu.cn

网络出版地址:http://kns.cnki.net/kcms/detail/50.1165.N.20191125.1034.042.html

能处理没有成对信息的数据。DeMeshNet 将去除网纹的过程分为两部分,先做网纹检测,再做图像修复。但是这两个模型都要求使用与带网纹人脸一一对应的网纹数据,然而在实际应用中要获得对应的网纹数据非常困难的。

将去网纹的过程转换为图像分离的过程如图 2 所示,提出分离对抗生成网络(Distangling-GAN)。该网络由分别代表带网纹人脸(X 域)、干净人脸(Y 域)和网纹(Z 域)的 3 个域,1 个生成器 G 以及 3 个判别器  $D_X, D_Y, D_Z$  组成。X 域图像经过分离网络 G, 得到 Y 域图像以及 Z 域图像;Y 域图像和 Z 域图像经过像素值叠加 L 合成 X 域图像。所以本文提出的网络的目标是训练一个分离图像的映射  $G: X \rightarrow (Y, Z)$ , 与此同时,训练 3 个判别器  $D_X, D_Y, D_Z$  与生成器 G 形成对抗,激励生成器能生成更好的图像。训练时主要通过极小化如下损失进行:

- 1) 前向循环损失:  $x \rightarrow G(x) = (\hat{y}, \hat{z}) \rightarrow (\hat{y} \oplus \hat{z}) = \hat{x} \approx x$ ;
- 2) 后向循环损失:  $(y, z) \rightarrow (y \oplus z) = \hat{x} \rightarrow G(\hat{x}) = (\hat{y}, \hat{z}) \approx (y, z)$ 。

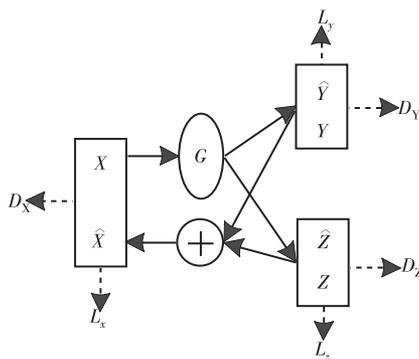


图 2 分离生成对抗网络框架

Fig. 2 The framework of disentangling-GAN

分离网络 G 采用自编码结构<sup>[16-17]</sup>,如图 3 所示。分离网络 G 由 1 个编码器和两个解码器组成,编码器将 X 域图像表示在特征空间,两个解码器将图像从特征空间分别生成 Y 域图像以及 Z 域图像。图像合成的部分采用像素值叠加。

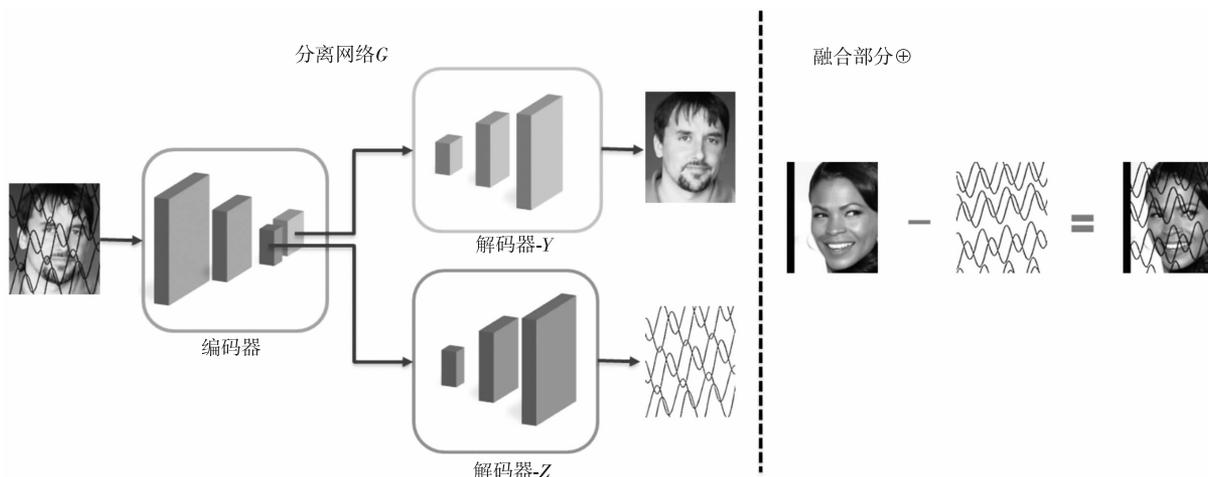


图 3 分离网络 G 与像素值叠加过程

Fig. 3 Separation of network G and pixel value superposed process

目前基于深度学习的针对网纹结构遮挡的图像修复方法<sup>[13-14,18]</sup>在训练时都需要使用对应的网纹数据输入模型中,但是在实际应用中很难能拿到训练所需要的网纹数据。普遍的解决方式是照着实际网纹数据的样子制作类似的网纹数据集作为训练集。但是网纹的制作过程中使用随机振幅、粗细、旋转角度以及贴网纹时的随机透明度的影响,使得难以制作出与实际网纹数据分布一致的网纹数据。这就导致了模型在实际使用的时候效果不佳,因为模型的效果非常依赖于制作的网纹数据与实际网纹数据的相近程度。

与现有的方法相比,本文提出的方法避免了输入与带网纹人脸一一对应的网纹数据,模型只需要输入带网

纹的人脸图像以及干净的人脸图像。在实际训练的过程中使用残差网纹代替了真实的网纹数据进行训练,一方面避免了对应网纹数据的输入,另一方面为深度学习模型加入了先验知识,提高模型的去网纹能力。实验结果表明,针对修复带网纹的人脸照片,图像修复后无论是纹理还是色差本文模型的结果与原图最为接近。

## 1 预备知识

本文的方法基于对抗生成网络以及图像修复领域,以下部分对对抗生成网络和图像修复做一个简短的介绍。

### 1.1 对抗生成网络

Goodfellow 等人<sup>[19]</sup>在 2014 年提出了对抗生成网络(GAN)使用对抗策略来学习一个生成模型。它由 1 个生成器以及 1 个判别器组成,在训练的时候极小化生成器损失,极大化判别器损失,交替地进行训练。之后由 Radford 等人<sup>[20]</sup>提出的 DCGAN 在对抗生成网络上做了一些改进,使得模型专门用于生成图像,但是模型生成图像的时候不受控制,不能生成指定的图像。更进一步地, Mirza 等人<sup>[21]</sup>提出了带有约束的对抗生成网络 cGAN,该网络可以根据标签生成指定的图像。为了利用不成对的数据进行训练, Cycle-GAN<sup>[15]</sup>, Dual-GAN<sup>[22]</sup>, Disco-GAN<sup>[23]</sup> 等被提出,它们能处理不成对的数据,让图片从一个域转化到另一个域进行风格迁移或图像转换。然而这些方法都没有考虑到遮挡的结构信息, DF-GAN<sup>[13]</sup> 考虑了遮挡的结构信息,但是训练的时候需要使用对应的网纹数据,网纹数据又往往不可能拿到。

### 1.2 图像修复

目前已有大量地方法用于图像修复,包括基于先验知识的方法<sup>[24-26]</sup>、使用复制粘贴的方法<sup>[27-29]</sup>、以及基于学习的方法<sup>[1, 9-10, 12-13, 18, 30]</sup>。一种早期的图像修复方法<sup>[26]</sup>使用图像的低秩结构,将图像修复问题转化为一个修复矩阵的优化问题,这对于较小且比较均匀的图像是有效的。另一方面,基于复制粘贴策略的方法效率很高,它搜索单张图片中最相近的补丁<sup>[27]</sup>或从额外的数据库中搜索<sup>[29]</sup>,但是当额外数据库中也找不到相似的补丁时该方法就失效了。近些年来,基于学习的方法,特别是基于深度学习的方法在图像修复领域产生了引人注目的效果。文献[31]将基于深度学习的图像修复方法总结为 3 类:第一类是基于卷积自编码结构的修复方法<sup>[1, 9, 11, 32]</sup>;第二类是基于生成对抗网络的修复方法<sup>[13, 20-23]</sup>;第三类是基于循环神经网络的修复方法<sup>[33-34]</sup>。第一类方法中,最初由 Phatak 等人<sup>[9]</sup>提出了上下文编码器,使用自编码结构并在预测的图像和原图之间使用  $L_2$  范数损失函数进行训练。文献[11]提出一种高分辨率的图像修复方法,该方法设计了一个学习图像内容的网络以及一个学习图像纹理的网络,具体以低分辨率图像修复的结果作为内容约束,并利用纹理约束的多尺度神经网络图像块合成方法进行高分辨率图像修复。文献[32]采用两个自编码结构网络来进行图像修复,第 1 个网络基于  $L_2$  距离对修复区域进行粗修复;第 2 个网络对第 1 个网络的结果进行纹理细节增强,获得了比较好的修复结果。针对人脸照片去网纹的应用中, DeMeshNet<sup>[14]</sup> 使用一个多任务的网络框架,先进行网纹检测,再进行图像修复。但是所有基于深度学习的图像修复方法在训练的时候要预先知道图像遮挡的位置,而在人脸照片去网纹这一应用中,网纹遮挡的信息是不能预先拿到的。

## 2 提出的方法

这一部分将详细介绍本文提出的方法 Distangling-GAN。如图 2 所示,该网络由带网纹人脸( $X$  域)、干净人脸( $Y$  域)、网纹( $Z$  域)等 3 个域,以及 1 个生成器  $G$  和 3 个判别器  $D_X, D_Y, D_Z$  组成。

$X$  域图像经过分离网络  $G$ , 得到  $Y$  域图像以及  $Z$  域图像;  $Y$  域图像和  $Z$  域图像经过像素值叠加  $\oplus$  合成  $X$  域图像。所以提出的网络的目标是训练 1 个分离图像的映射  $G: X \rightarrow (Y, Z)$ , 与此同时, 训练 3 个判别器  $D_X, D_Y, D_Z$  与生成器  $G$  形成对抗, 激励生成器能生成更好的图像。其中训练采用的  $Z$  域的网纹数据通过  $X$  域和  $Y$  域数据的像素值差得到, 也就是说该模型不需要使用网纹数据。

### 2.1 分离网络

模型的分离网络  $G$  的目标是将  $X$  域数据中的网纹和干净人脸图片分开, 分成  $Y$  域数据(干净人脸)以及  $Z$  域数据(网纹), 即分离网络是这样映射:  $G: X \rightarrow (Y, Z)$ 。在图 3 中, 分离网络采用自编码结构<sup>[8, 15-16]</sup>。首先使用一个编码器将  $X$  域中的图像  $x$  映射到特征潜在表示层(Latent representation)  $G_{\text{Enc}}(x)$ , 在潜在表示层将  $G_{\text{Enc}}(x)$  分解成两部分  $G_{\text{Enc}|Y}(x)$  和  $G_{\text{Enc}|Z}(x)$ , 分别对应  $Y$  域数据的特征层和  $Z$  域数据的特征层; 然后使用两个解码器  $G_{\text{Dec-}Y}, G_{\text{Dec-}Z}$  将特征层数据转换成干净的人脸数据  $\hat{y}$  及网纹数据  $\hat{z}$ 。分离过程可用如下公式表示:

$$\begin{cases} \hat{y} = G_{\text{Dec}-Y}(G_{\text{Enc}|Y}(x)) \\ \hat{z} = G_{\text{Dec}-Z}(G_{\text{Enc}|Z}(x)) \end{cases} \quad (1)$$

## 2.2 替代网纹及融合部分

目前基于深度学习的针对网纹结构遮挡的图像修复方法<sup>[13-14,18]</sup>中都需要使用对应的网纹数据作为训练集输入模型中,但是在实际应用中很难能拿到训练所需要的网纹数据,于是人们就对比网纹遮挡的样子自己制作类似的网纹数据集作为训练集。但是网纹的制作过程中使用随机的振幅、粗细、旋转角度以及贴网纹时的随机透明度,使得难以制作出与目标网纹分布一致的网纹数据。这就导致了模型在实际使用的时候效果不佳,因为模型的效果非常依赖于网纹数据制作的好坏。

为了避免对应网纹数据的制作,提出的方法使用残差网纹来代替真实的网纹数据。如图 4 所示,使用 Y 域数据与 X 域数据做像素值差,经过转换后得到模型训练使用的网纹数据。

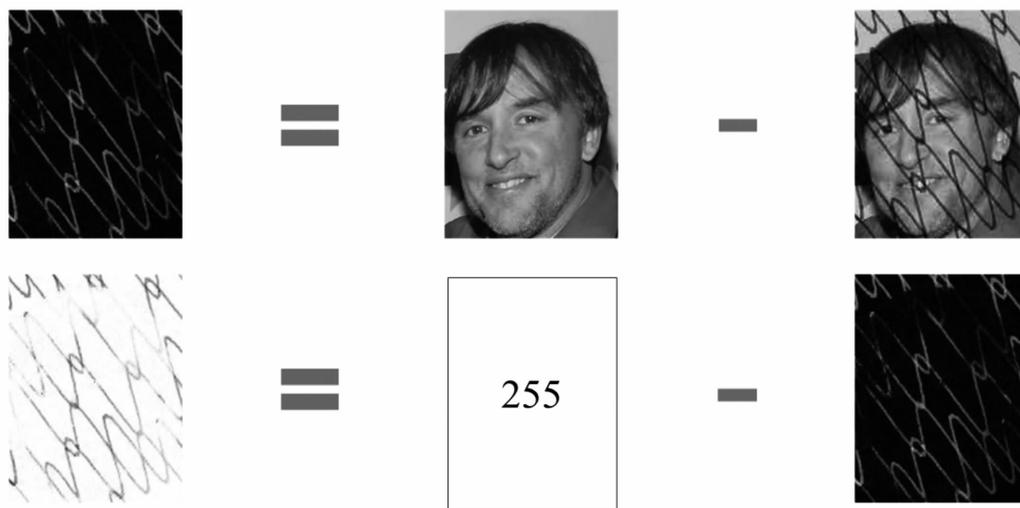


图 4 残差网纹的生成

Fig. 4 Residual texture generation

图 4 中的 255 代表一张像素值全为 255 的白色图像。也正是由于模型的网纹是利用图像像素值差得到的,所以模型中干净人脸与网纹融合的部分不必使用神经网络去学习。与 DF-GAN<sup>[12]</sup>相比,该模型直接采用像素值叠加融合图像(图 3),这样比 DF-GAN 的网络大小几乎小了一半,并且作为分离网络的逆过程更能监督分离网络的学习。

## 2.3 判别网络

整个模型需要输入 3 部分数据:  $\{x_i\}_{i=1}^N \in X$ : X 域数据,即带网纹人脸图;  $\{y_i\}_{i=1}^N \in Y$ : Y 域数据,即干净人脸图;  $\{z_i\}_{i=1}^N \in Z$ : Z 域数据,即残差网纹图。模型也会通过分离结构以及合成结构生成这 3 个域的图像:  $(\hat{y}, \hat{z}) = G(x)$ ,  $\hat{x} = y \oplus z$ 。所以在生成图片过后,需要 3 个判别器来判断图片是来自于真实的数据,还是由生成器生成的数据,以形成对抗训练。3 个域的判别器使用相同的结构,每个判别器由 5 个卷积层构成,前 4 层采用  $4 \times 4$  的卷积核以及 leakyrelu 激活方式,最后一层使用 sigmoid 激活函数。

## 2.4 对抗训练

对抗训练是采用生成器与判别器之间的一个极小极大化策略,目标函数如下:

$$\min_G \max_{D_X, D_Y, D_Z} L_{\text{GAN}}(G, D_X, D_Y, D_Z) \quad (2)$$

其中:

$$\begin{aligned} L_{\text{GAN}}(G, D_X, D_Y, D_Z) = & E_{x \sim P_{\text{data}}(x)} [\ln(1 - D_Y(\hat{y})) + \ln(1 - D_Z(\hat{z}))] + \\ & E_{y \sim P_{\text{data}}(y), z \sim P_{\text{data}}(z)} [\ln(1 - D_X(y \oplus z))] + E_{x \sim P_{\text{data}}(x)} [\ln D_X(x)] + \\ & E_{y \sim P_{\text{data}}(y)} [\ln D_Y(y)] + E_{z \sim P_{\text{data}}(z)} [\ln D_Z(z)]. \end{aligned} \quad (3)$$

$P_{\text{data}}(x), P_{\text{data}}(y), P_{\text{data}}(z)$  分别代表 3 个域中真实的数据分布,  $\hat{y}, \hat{z}$  代表由生成器  $G$  生成的数据(见(1)式),  $\hat{x}$  代表

由 Y 域与 Z 域合成的数据,即  $\hat{x} = y \oplus z$ 。

在模型中分离网络 G 与确定性融合部分组成了一个循环,所以参考 Cycle-GAN<sup>[15]</sup> 的损失函数设置。其中,  $\lambda_1, \lambda_2$  是权重参数。

$$L_{\text{cyc}}(G) = \lambda_1 E_{x \sim P_{\text{data}}(x)} \|F(G(x)) - x\| + \lambda_2 E_{y \sim P_{\text{data}}(y), z \sim P_{\text{data}}(z)} \|G(y \oplus z) - (y, z)\|。 \quad (4)$$

综上所述,最后要优化的损失函数如下:

$$\min_G \max_{D_x, D_y, D_z} L_{\text{GAN}}(G, D_x, D_y, D_z) + L_{\text{cyc}}(G)。 \quad (5)$$

在实际求解的时候,按照对抗训练的算法框架<sup>[19]</sup> 并且参考 Cycle-GAN<sup>[15]</sup> 的求解过程,将上述损失函数分成以下几个损失函数,将上述的极小极大化的目标函数换成 6 个极小化损失函数:

$$Y \text{ 域生成损失: } G_{\text{loss}Y} = E_{x \sim P_{\text{data}}(x)} [-\ln(D(\hat{y}))] + L_{\text{cyc}}(G);$$

$$Y \text{ 域判别损失: } D_{\text{loss}Y} = E_{y \sim P_{\text{data}}(y)} [\ln D_Y(y)] + E_{x \sim P_{\text{data}}(x)} [\ln(1 - D(\hat{y}))];$$

$$Z \text{ 域生成损失: } G_{\text{loss}Z} = E_{x \sim P_{\text{data}}(x)} [-\ln(D(\hat{z}))] + L_{\text{cyc}}(G);$$

$$Z \text{ 域判别损失: } D_{\text{loss}Z} = E_{z \sim P_{\text{data}}(z)} [\ln D_Z(z)] + E_{x \sim P_{\text{data}}(x)} [\ln(1 - D(\hat{z}))];$$

$$X \text{ 域生成损失: } G_{\text{loss}X} = E_{x \sim P_{\text{data}}(x)} [-\ln(D(\hat{x}))] + L_{\text{cyc}}(G);$$

$$X \text{ 域判别损失: } D_{\text{loss}X} = E_{x \sim P_{\text{data}}(x)} [\ln D_X(x)] + E_{y \sim P_{\text{data}}(y)} E_{z \sim P_{\text{data}}(z)} [\ln(1 - D(\hat{x}))]。$$

然后对于输入的每一批次数据  $\{x^{(1)}, x^{(2)}, \dots, x^{(m)}\}, \{y^{(1)}, y^{(2)}, \dots, y^{(m)}\}, \{z^{(1)}, z^{(2)}, \dots, z^{(m)}\}$ , 利用随机梯度法依次更新生成网络和判别网络:

$$\begin{cases} \nabla_{\theta} \frac{1}{m} \sum_{i=1}^m G_{\text{loss}Y}(x^{(i)}, y^{(i)}, z^{(i)}), \nabla_{\theta} \frac{1}{m} \sum_{i=1}^m D_{\text{loss}Y}(x^{(i)}, y^{(i)}, z^{(i)}) \\ \nabla_{\theta} \frac{1}{m} \sum_{i=1}^m G_{\text{loss}Z}(x^{(i)}, y^{(i)}, z^{(i)}), \nabla_{\theta} \frac{1}{m} \sum_{i=1}^m D_{\text{loss}Z}(x^{(i)}, y^{(i)}, z^{(i)}), \\ \nabla_{\theta} \frac{1}{m} \sum_{i=1}^m G_{\text{loss}X}(x^{(i)}, y^{(i)}, z^{(i)}), \nabla_{\theta} \frac{1}{m} \sum_{i=1}^m D_{\text{loss}X}(x^{(i)}, y^{(i)}, z^{(i)}) \end{cases} \quad (6)$$

直至达到终止条件。

### 3 实验分析

为了说明本研究工作的有效性,本文在 4 个公开人脸数据集上与其他多种深度学习的方法进行了比较。实验采用的操作系统为 Ubuntu 16.04, 编程语言为 Python 3.6, 使用 Tensorflow 框架 1.1 版本, 系统处理器为 Xeon E5, 64 GB 内存, 配有一块 NVIDIA Taitan XP 显卡, 所有模型训练时迭代 20 次全数据集。为了进行客观验证, 图像修复后采用峰值信噪比 (PSNR) 和结构相似性 (SSIM)<sup>[31]</sup> 对实验结果进行测评。

#### 3.1 数据集准备

The AR face database<sup>[36]</sup>: AR 人脸数据集由 126 个人 (70 名男性, 56 名女性) 超过 4 000 张人脸图片组成。这些照片是具有不同面部表情、照明条件和遮挡 (太阳镜和围巾) 的正面视图。在本文的实验中选取了没有遮挡的图片 2 600 张, 其中 1 820 张作为训练集, 剩下的 780 张作为测试集。

The CMU MultiPIE face database<sup>[37]</sup>: 该数据集由 337 人在 15 个观察点和 19 个照明条件下拍照, 同时包含了一系列面部表情, 总共超过 750 000 张图像。本文选取了其中人脸正面视图以及光照强度适中的图片进行实验, 一共选出 8 089 张图片用作训练, 3 465 张图片用作测试。

The Color FERET face database<sup>[38]</sup>: 在带条件拍摄的人脸图像数据集中还选择了 FERET 数据集进行实验。该数据集共有 11 338 张人脸图片, 包括了 994 个在不同年龄段的人, 分 13 个姿势拍出的照片。在实验中, 本文选取了标记为 'fa' (正面照) 和 'fb' (与正面角度类似) 这两个姿势的图片, 共 980 张图片作为训练集, 420 张图片作为测试集。

The LFW face database<sup>[6]</sup>: 除此之外, 本文也在无条件的人脸图像数据集上对模型进行了测试验证。LFW (Labeled Faces in the Wild) 数据集从网上搜集了超过 1 680 个人, 13 000 张人脸图像。经过筛选, 选出 6 950 张图像作为训练集, 2 980 张图像作为测试集。

网纹数据制作: 由于没有大规模的网纹照片数据集, 所以根据文献<sup>[18]</sup> 中的方法来制作网纹数据。第 1 步,

通过对正弦、余弦函数的振幅、相位、旋转角度以及线条粗细进行随机扰动获得网纹图片,如图 5 所示。

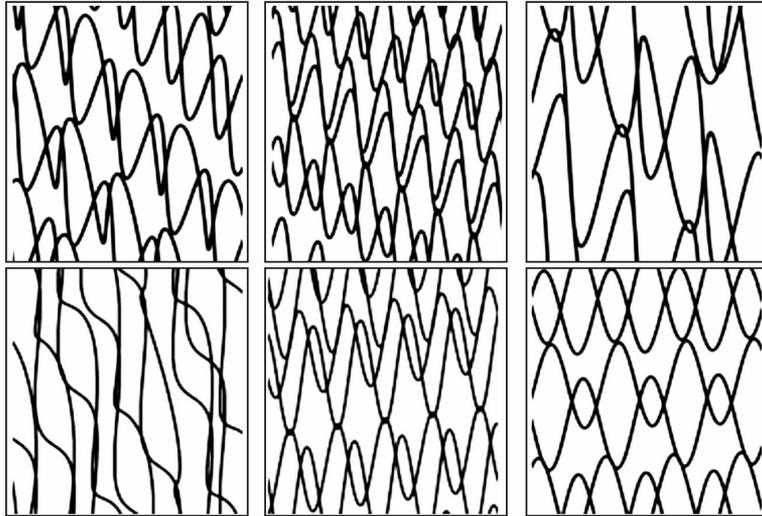


图 5 生成的网纹图

Fig. 5 Generated mesh net

第 2 步,根据(7)式将网纹与人脸照片合成。其中, $f_{\text{net}}$ 代表网纹图像, $y$ 代表人脸图像, $x$ 代表合成后的带网纹人脸图。 $\alpha \in [0, 1]$ 控制着图像合成时网纹的透明度, $\alpha$ 越大网纹的颜色越深。实际合成时,在区间 $[0.1, 0.9]$ 按均匀分布随机选取一个 $\alpha$ 作为合成透明度。制作好了带网纹人脸图后就可以根据前面残差网纹的制作方法做出残差网纹了。

$$x = \begin{cases} \alpha f_{\text{net}} + (1 - \alpha)y, & f_{\text{net}} < 1 \\ y, & f_{\text{net}} = 1 \end{cases}, \quad (7)$$

数据预处理与模型参数设置:上述数据集的所有人脸图像都根据图像人眼位置将图片裁剪成 $128 \times 128$ 大小。模型训练采用 ADAM 求解器,学习率设置为 0.000 2,批次大小设为 1,损失函数中的权重参数取 $\lambda_1 = 10$ , $\lambda_2 = 10$ ,分离网络中图像的潜在表示层尺寸为 $256 \times 32 \times 32 \times 1$ ,其中, $255 \times 32 \times 32 \times 1$ 作为干净人脸的特征表示,在经过 5 个残差块以及解码器后获得干净人脸图片; $1 \times 32 \times 32 \times 1$ 作为网纹的特征表示,经过 1 个残差块以及解码器后获得网纹图片。

### 3.2 模型结果

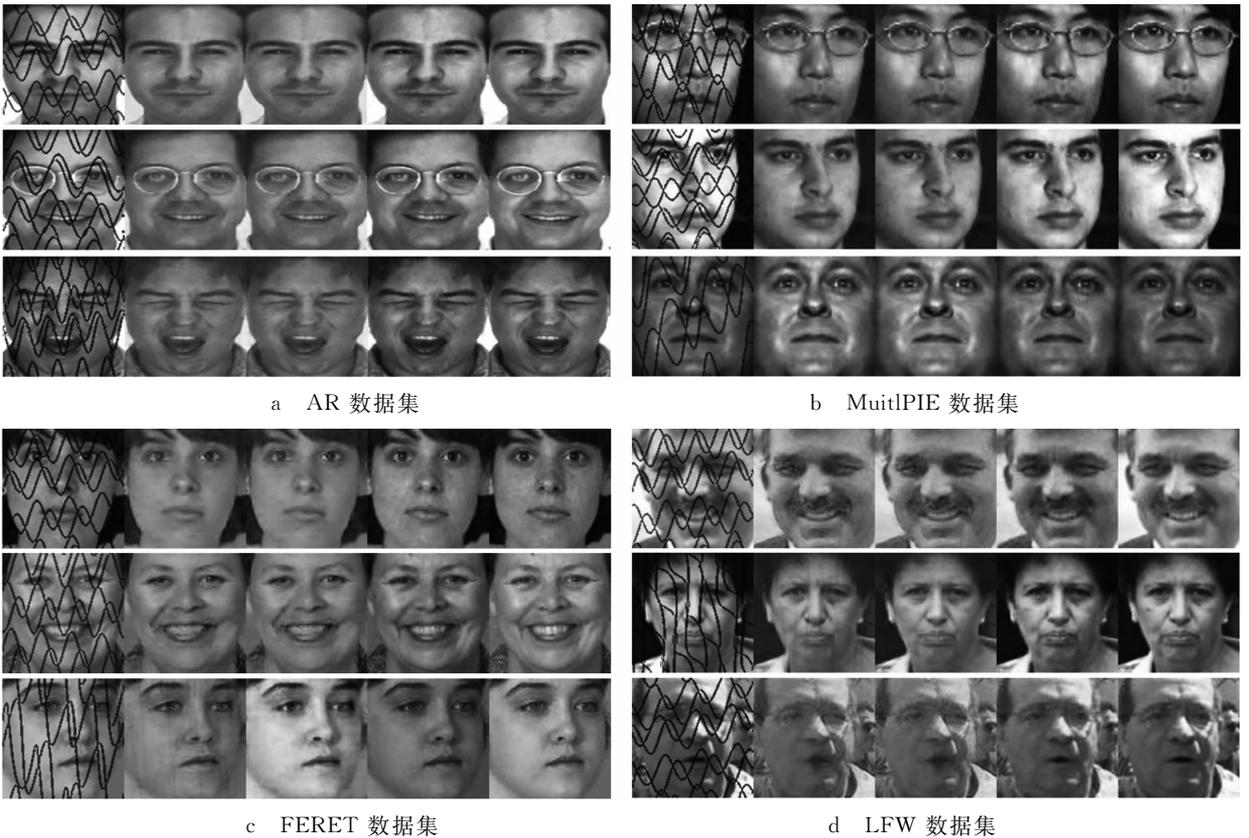
从指标结果上看,在大部分数据集上本文的方法都明显好于另外两种方法(表 1),只有在 MultiPIE 数据集上比 DF-GAN 差,但是差距不大。图 6 中,从视觉效果上看,CycleGAN 能去除大部分网纹但是有些还能看到网纹痕迹,DF-GAN 去除网纹过后几乎看不到网纹的痕迹,但是修复后的图片有些模糊纹理信息消失,与真实图片相比存在色差。无论是纹理还是色差,图像修复后本文模型的结果与原图最为接近。

表 1 AR, MultiPIE, FERET, LFW 数据集上各模型的指标结果

Tab. 1 Index results for each model on the AR, MultiPIE, FERET, LFW dataset

方法	峰值信噪比(PSNR)				结构相似性(SSIM)			
	AR	CMU_PIE	FERET	LFW	AR	CMU_PIE	FERET	LFW
带网纹图	15.66	18.92	18.56	18.23	0.451 8	0.531 6	0.433 6	0.450 4
Cycle-GAN	21.72	23.13	25.72	23.08	0.875 7	0.822 7	0.861 1	0.832 9
DF-GAN	23.85	28.21*	28.15	23.18	0.916 8	0.917 6*	0.931 0	0.869 0
本文方法	31.76*	26.27	29.12*	25.25*	0.936 3*	0.864 7	0.948 1*	0.925 0*

注:\*表示同数据集下指标表现最优的算法。



注:各小图每行从左至右依次为带网纹人脸图、CycleGAN 结果图、DF-GAN 结果图、本文结果图和原图。

图 6 不同方法在各数据集上的修复结果

Fig. 6 Repair results of different methods on each data set

## 4 结束语

本文提出了一种基于对抗生成网络的针对网纹遮挡的人脸图像修复技术。将恢复干净人脸图像看成是一个图像分离的过程,使用 1 个分离网络与 3 个判别网络构建一个对抗生成网络来模拟整个过程。整个网络设置 3 个域(带网纹图像、干净人脸图像、网纹),将图像分离的过程看作是域与域之间互相转化的过程。在分离网络中采用了自编码结构来生成图像,融合的部分采用图像像素值做差,分离与融合构成一个循环。该模型不使用对应的网纹数据作为输入,并且实验结果表明了提出的方法对于人脸图像在消除网纹方面的有效性。

注:本文使用的真人图像中,图 1 已得到图像本人授权,且仅适用于本文,图 3、图 4 中的真人图像均来源于公开的人脸数据集,具体可见 <http://mmlab.ie.cuhk.edu.hk/projects/CelebA.html>。

## 参考文献:

- [1] LI Y, LIU S, YANG J, et al. Generative face completion [C]//Conference on Computer Vision and Pattern Recognition (CVPR). Hawaii, USA:IEEE, 2017.
- [2] BENGIO Y, COURVILLE A, VINCENT P. Representation learning: a review and new perspectives[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2013, 35(8):1798-1828.
- [3] 张玉宏, 秦志光. 深度学习的方法论辨析[J]. 重庆理工大学学报(社会科学), 2018, 32(6):20-26.  
ZHANG Y H, QIN Z G. Analysis of methodology of deep learning[J]. Journal of Chongqing University of Technology (Social Science), 2018, 32(6):20-26.
- [4] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. Imagenet classification with deep convolutional neural networks[C]//Proceedings of the 25th International Conference on Neural Information Processing Systems. Lake Tahoe, USA:MIT Press, 2012.
- [5] TAIGMAN Y, YANG M, RANZATO M, et al. DeepFace: closing the gap to human-level performance in face verification[C]//Conference on Computer Vision and Pattern Rec-

- ognition (CVPR). Columbus, USA; IEEE, 2014.
- [6] HUANG G B, MATTAR M, BERG T, et al. Labeled faces in the wild: a database for studying face recognition in unconstrained environments[EB/OL]. (2008-09-16) [2019-11-25]. <https://hal.inria.fr/inria-00321923>.
- [7] ZHANG Z, LUO P, LOY C C, et al. Facial landmark detection by deep multi-task learning[C]//European Conference on Computer Vision. Zurich. Switzerland; Springer, 2014.
- [8] SCHROFF F, KALENICHENKO D, PHILBIN J. FaceNet: a unified embedding for face recognition and clustering [C]//2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Boston, USA; IEEE, 2015.
- [9] PATHAK D, KRAHENBUHL P, DONAHUE J, et al. Context encoders: feature learning by inpainting[C]//IEEE Conference on Computer Vision & Pattern Recognition (CVPR). Las Vegas, USA; IEEE, 2016.
- [10] YE H R A, CHEN C, LIM T Y, et al. Semantic image inpainting with deep generative models [C]//IEEE Conference on Computer Vision & Pattern Recognition (CVPR). Hawaii, USA; IEEE, 2017.
- [11] YANG C, LU X, LIN Z, et al. High-resolution image inpainting using multi-scale neural patch synthesis [C]//IEEE Conference on Computer Vision & Pattern Recognition (CVPR). Hawaii, USA; IEEE, 2017.
- [12] HUANG H, HE R, SUN Z, et al. Wavelet-SRNet: a wavelet-based CNN for multi-scale face super resolution[C]//2017 IEEE International Conference on Computer Vision (ICCV). Venice, Italy; IEEE, 2017.
- [13] LI Z. H, HU Y B, HE R. Learning disentangling and fusing networks for face completion under structured occlusions[EB/OL]. (2017-12-13) [2019-11-25]. <https://arxiv.org/pdf/1712.04646.pdf>.
- [14] ZHANG S, HE R, SUN Z, et al. DeMeshNet: blind face inpainting for deep meshface verification[J]. IEEE Transactions on Information Forensics and Security, 2018, 13 (3): 637-647.
- [15] ZHU J Y, PARK T, ISOLA P, et al. Unpaired image-to-image translation using cycle-consistent adversarial networks[C]//2017 IEEE International Conference on Computer Vision (ICCV). Venice, Italy; IEEE, 2017.
- [16] BENGIO Y, LAMBLIN P, POPOVICI D, et al. Greedy layer-wise training of deep networks[C]//Proceedings of the 19th International Conference on Neural Information Processing Systems. Vancouver, Canada; MIT Press, 2007.
- [17] VINCENT P, LAROCHELLE H, BENGIO Y, et al. Extracting and composing robust features with denoising autoencoders [C]//International Conference on Machine Learning. Helsinki; ACM, 2008.
- [18] ZHANG S, HE R, SUN Z, et al. Multi-task ConvNet for blind face inpainting with application to face verification [C]//International Conference on Biometrics. Halmstad, Sweden; IEEE, 2016.
- [19] GOODFELLOW I J, POUGET-ABADIE J, MIRZA M, et al. Generative adversarial nets[C]//International Conference on Neural Information Processing Systems. Montreal, Canada; MIT Press, 2014.
- [20] CHINTALA S, RADFORD A, METZ L. Unsupervised representation learning with deep convolutional generative adversarial networks [EB/OL]. (2017-12-13) [2019-06-28]. <https://arxiv.org/pdf/1511.06434.pdf>.
- [21] OSINDERO S, MIRZA M. Conditional generative adversarial nets[EB/OL]. (2017-12-13) [2019-06-28]. <https://arxiv.org/pdf/1411.1784.pdf>.
- [22] YI Z L, ZHANG H, TAN P, et al. DualGAN: unsupervised dual learning for image-to-image translation [C]//2017 IEEE International Conference on Computer Vision (ICCV). Venice, Italy; IEEE Press, 2017.
- [23] KIM T, CHA M, KIM H, et al. Learning to discover cross-domain relations with generative adversarial networks[C]//International Conference on Machine Learning. Sydney, Australia; ACM, 2017.
- [24] AFONSO M V, BIOUCAS-DIAS J M, FIGUEIREDO M A T. An augmented lagrangian approach to the constrained optimization formulation of imaging inverse problems [J]. IEEE Transactions on Image Processing, 2011, 20 (3): 681-695.
- [25] HE R, ZHENG W S, TAN T N, et al. Half-quadratic-based iterative minimization for robust sparse representation[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2014, 36(2): 261-275.
- [26] HU Y, ZHANG D, YE J, et al. Fast and accurate matrix completion via truncated nuclear norm regularization[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2013, 35(9): 2117-2130.
- [27] BARNES C, SHECHTMAN E, FINKELSTEIN A, et al. PatchMatch: a randomized correspondence algorithm for structural image editing [J/OL]. ACM Transactions on Graphics, 2009, 28(3): No. 24. (2009-07-27) [2019-11-25]. <https://dl.acm.org.citation.cfm?doid=1531326.1531330>.
- [28] EFROS A A, LEUNG T K. Texture synthesis by non-parametric sampling [C]//2002 IEEE International Conference on Computer Vision (ICCV). Beijing; IEEE, 2002.
- [29] HAYS J, EFROS A A. Scene completion using millions of photographs[J]. ACM Transactions on Graphics, 2008, 26 (3): 4-11.
- [30] LIU S F, PAN J S, YANG M H. Learning recursive filters for low-level vision via a hybrid neural network[C]//European Conference on Computer Vision (ECCV). Amster-

dam; Springer, 2016.

[31] 强振平, 何丽波, 陈旭, 等. 深度学习图像修复方法综述[J]. 中国图象图形学报, 2019, 24(3): 127-143.

QING Z P, HE L B, CHEN X, et al. Survey on deep learning image inpainting methods[J]. Journal of Image and Graphics, 2019, 24(3): 447-463.

[32] YU J H, LIN Z, YANG J M, et al. Generative image inpainting with contextual attention [C]//Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Salt Lake City, USA: IEEE, 2018.

[33] OORD A V D, KALCHBRENNER N, KAVUKCUOGLU K. Pixel recurrent neural networks [C]//Proceedings of the 33rd International Conference on Machine Learning. New York, USA: JMLR, 2016.

[34] BURLIN C, LE C Y, DUPERIER L. Deep image inpaint-

ing[EB/OL]. (2017-07-02)[2019-06-28]. <http://cs231n.stanford.edu/reports/2017/pdfs/328.pdf>.

[35] WANG Z, BOVIK A C, SHEIKH H R, et al. Image quality assessment: from error visibility to structural similarity [J]. IEEE Transactions on Image Processing, 2004, 13(4): 600-612.

[36] BENAVENTE R., MARTINEZ A. The AR face database [R]. CVC technical report, 1998.

[37] GROSS R, MATTHEWS I, COHN J, et al. Multi-pie [J]. Image and Vision Computing, 2010, 28(5): 807-813.

[38] PHILLIPS P J, MOON H, RIZVI S A, et al. The FERET evaluation methodology for face-recognition algorithms [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2000, 22(10): 1090-1104.

## Method for Eliminating Mesh of Face Photograph Based on Generative Adversarial Networks

ZHANG Ke<sup>1,3</sup>, BAI Fusheng<sup>1,3</sup>, WU Zhiyou<sup>1,3</sup>, PI Jiatian<sup>2,3</sup>, ZHAO Lijun<sup>3,4</sup>

(1. School of Mathematical Sciences, Chongqing Normal University;

2. College of Computer & Information Science, Chongqing Normal University, Chongqing 401331;

3. Intelligent Finance & Large Data Analysis, With Key Laboratory in Chongqing, Chongqing 401331;

4. Mashang Consumer Finance Co, Ltd, Chongqing 401331, China)

**Abstract:** [Purposes] In the field of image restoration, it is very important to repair the face images with mesh occlusion for improving the accuracy of face verification. [Methods] At present, the model based on deep learning for repairing mesh face requires input of corresponding mesh data in training, but it is very difficult to obtain the corresponding mesh data in practical application. In order to avoid using the corresponding mesh data to effectively repair the face image, a disentangling generative adversarial net is proposed. The residual mesh used to replace the real mesh is generated by using the pixel difference between the occluded face photograph and the clean face photograph. The disentangling generative adversarial net consists of a disentangling network and three discrimination networks for image restoration. [Findings] The experimental results show that the proposed method is effective in eliminating the reticulation of face images. [Conclusions] For the occlusion of the mesh structure of the face image, a good repair effect can still be achieved without using the corresponding mesh data.

**Keywords:** generative adversarial net; face completion; eliminating mesh

(责任编辑 许 甲)