

热点追踪

抗姿态和遮挡的人脸生成与识别

重庆大学 段青言 张磊

姿态变化和面部遮挡是影响人脸识别最主要的两个因素。对于姿态变化，通常采用抗姿态的特征表达和基于生成对抗网络 (Generative Adversarial Net, GAN) 的人脸正面化这两种方式进行解决。对于面部遮挡，基于 GAN 模型的人脸修复方法也层出不穷，这些方法更多地关注正面或近正面的人脸面部结构以及像素细节，而非身份判别性。可见，姿态变化和面部遮挡通常被当作两个单独的任务来分别加以解决。然而，在实际生活中，这两种情况常常同时发生，且逐渐演变成为一种富有挑战且有待研究的问题。如图 1 中，第一行和第二行的输入图片姿态角分别为 45° 和 60° ，其中，(a) 为输入的侧面遮挡人脸，(b) 为我们提出的 BoostGAN 的生成结果，(c) 和 (d) 分别为现有两种人脸正面化方法的生成结果，(e) 为真实的正面人脸。从图 1 可以看出，当侧面人脸出现部分遮挡时，(c) 和 (d) 作为仅针对姿态变化的人脸正面化方法出现了不同程度的生成误差。

解决姿态变化和面部遮挡的混合问题，直接的想法是分两步进行处理，即先采用人脸修复方法去遮挡，然后人脸正面化。然而，两步法容易产生较大的误差且依

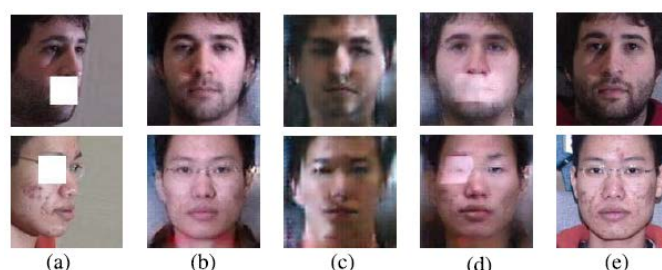


图 1 遮挡对现有有人脸正面化方法的影响

赖大量训练样本。我们提出的 BoostGAN 是一种端到端的集成式生成模型，如图 2 所示，采用多张局部遮挡的图片作为模型的输入，来完备身份和纹理信息。该网络由一个深度的编-解码器 (即粗糙网络) 和一个浅层的集成网络 (即精细网络) 组成。粗糙网络用于在多重遮挡和大姿态变化的人脸图像上实现粗糙的正面化和去遮挡生成。而精细网络旨在通过集成多个中间输出的互补信息，生成干净、正面的人脸图像并保持身份特异性。

更进一步，考虑到噪声作为先验知识被广泛应用至图像修复中，以及人脸修复和人脸正面化两个任务之间的协同作用，我们提出一种遮挡掩模引导下的两阶段生成对抗网络 (TSGAN)，如图 3 所示。该网络主要包含



图 2 端到端由粗糙到精细生成的 BoostGAN 框架

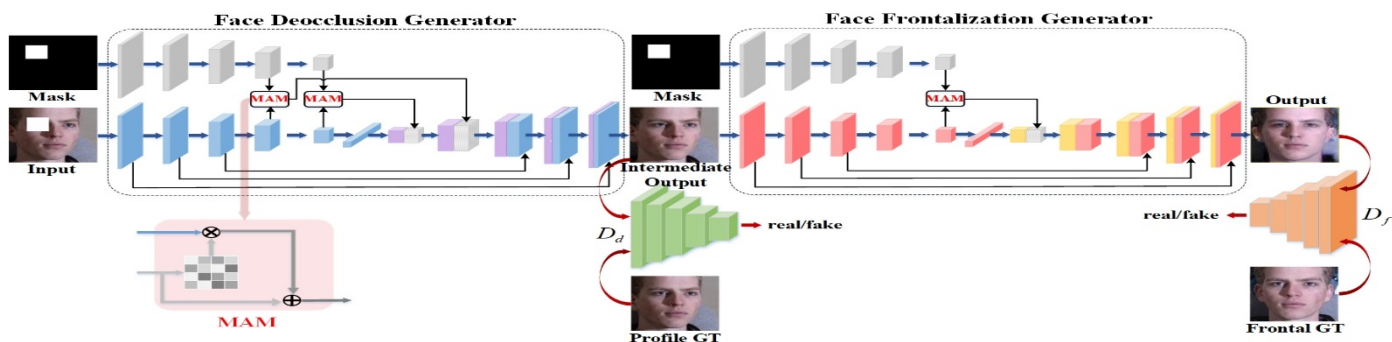


图3 两阶段生成对抗网络 (Two-Stage GAN, TSGAN)

3 个模块，即人脸去遮挡模块、人脸正面化模块和掩模注意力模块 (Mask Attention Module, MAM)。前两个模块分别被设计用于不同的阶段，而 MAM 则在两个阶段中均被部署。在第一个阶段，作为一种重要的先验知识，引入遮挡掩模来拟合输入图像中的噪声，作为辅助信号帮助 TSGAN 完成人脸修复。MAM 使得人脸去遮挡模块更多的关注和更好地填充侧面人脸图像上的“空洞”，如图 4 所示。在第二阶段，由第一阶段生成的无遮挡侧面人脸作为人脸正面化模块的输入，通过 MAM 进一步获得最终逼真的正面图像。值得注意的是，TSGAN 模型是一个端到端的结构。此外，为了更有效地分别监督两个阶段保持身份的一致性和提高身份相关特征的判别性，提出针对去遮挡和正面化的双重三元损失来联合训练 TSGAN 的两个阶段。

在约束和非约束的人脸图像数据集上，定量和定性

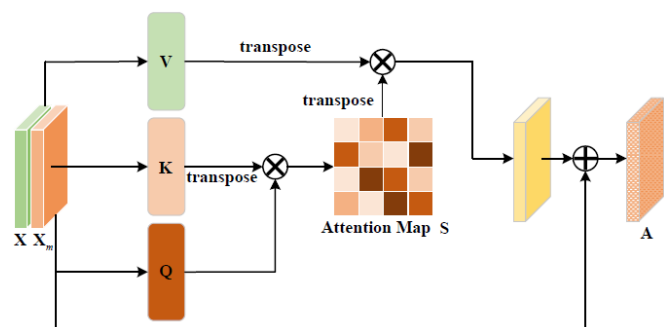


图4 遮掩模注意力模块 (MAM)

实验表明了提出的两个模型 BoostGAN 和 TSGAN 对遮挡、侧面人脸生成和识别的优越性，达到 SoTA。

以上 2 个成果分别被国际期刊 IEEE Transactions on Neural Networks and Learning Systems (2020) 和 IEEE Transactions on Circuits and Systems for Video Technology (2021)接收。

责任编辑 储珺



段青言

2021 年 6 月博士毕业于重庆大学，现为重庆邮电大学讲师，主要研究方向为人脸识别、人脸生成、深度学习。

Email: duanqy@cqupt.edu.cn



张磊

重庆大学教授，博士生导师，IEEE/CCF 高级会员。主要研究方向为开放环境视觉感知、深度学习、迁移学习等。

Email: leizhang@cqu.edu.cn