



# 人脸图像智能化处理研究

自然科学类学术论文

## • 01 行业背景 ——响应科技兴警，助力智慧公安



习近平总书记在全国公安会议强调，“当前犯罪形势发生了新变化，给公安工作提出了新挑战。但**科技兴警**的根本战略不能变，**依靠科技手段打击犯罪的思路不能动摇。**”

近年来，随着摄像和监控设备的普及，**犯罪图像证据**年均增幅达**34%**以上，**人脸图像智能化处理**成为破案关键。

刑侦犯罪图像数据（2010-2022）



应用场景

口罩去除



年龄变化



人像清晰



肤色处理

## • 02 研究内容

### ——基于生成对抗网络和注意力机制的人脸图像修复技术研究

- 卷积神经网络CNN
- 生成对抗网络GAN
- 残差网络ReNnet
- 组成三级网络**CLGN**

深度学习方法



图像修复技术



目标：根据人脸图像中的已知部分修复未知区域

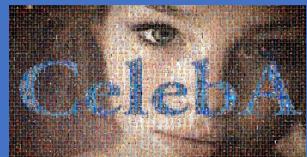
大数据训练



- 训练**200\*10**批次
- 全程耗时**1150**小时
- GPU: GeForce RTX 3090 Ti (31GB)



用大数据人脸图像集训练神经网络



采集**202, 599** 个彩色人脸图像，**10, 177** 个身份，数据总量超**200K 2000张**不同覆盖了大小的不规则掩膜 (QD-IMD) <https://github.com/karfly/qd-imd>

## • 03 技术痛点

生成**整体结构复杂**和**细节高保真**的人脸图像内容仍是一个巨大的挑战

### 传统图像修复算法

### 生成式图像修复模型

类别

基于扩散的修复算法  
(diffusion-based)



基于纹理的修复算法  
(patch-based)



CE模型  
(Context Encoder)



U-Net 网络  
(Ronneberger , 2015 )



存在问题

无法处理**背景复杂**, **缺损区域较大**的  
图像

当修复**不规则区域**破损的图像时, 仍会出  
现失真模糊现象

修复后存在伪影、结构扭曲、纹理模糊

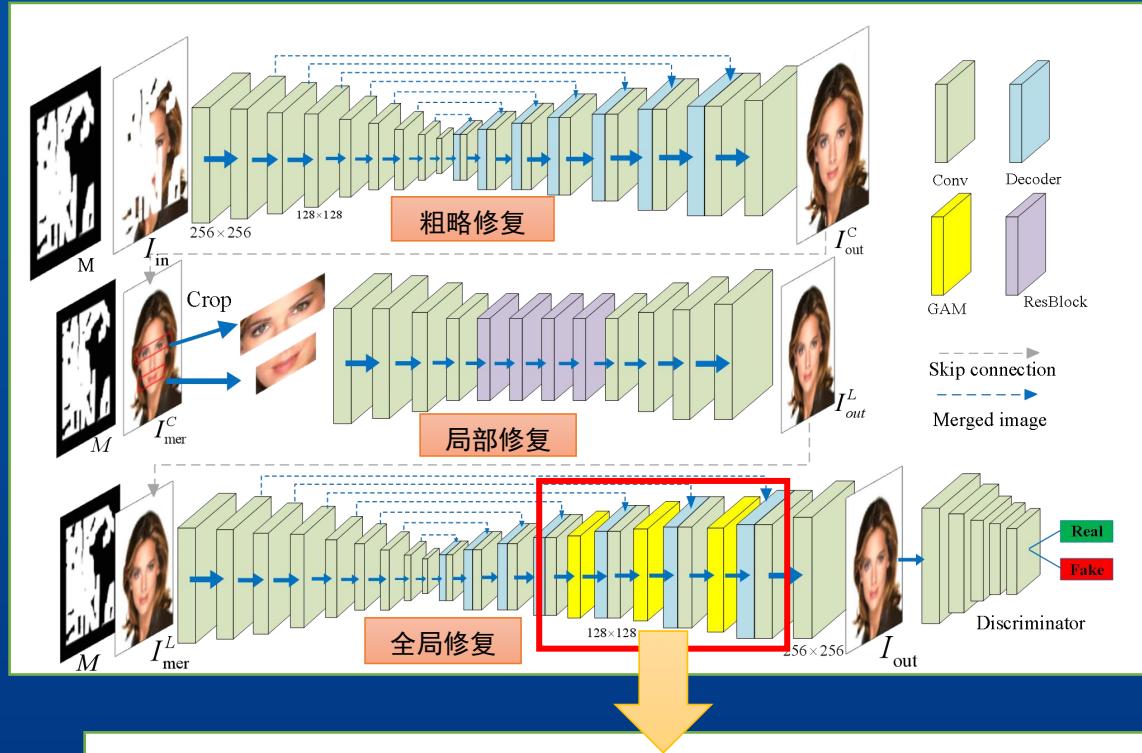
**技术痛点:** 无法准确传播**全局上下文信息**  
无法修复大面积空洞

无法合成真实**局部细节**,  
有效感受野**小**, 特征依赖**高**



返璞归真  
FAN PU GUI ZHEN

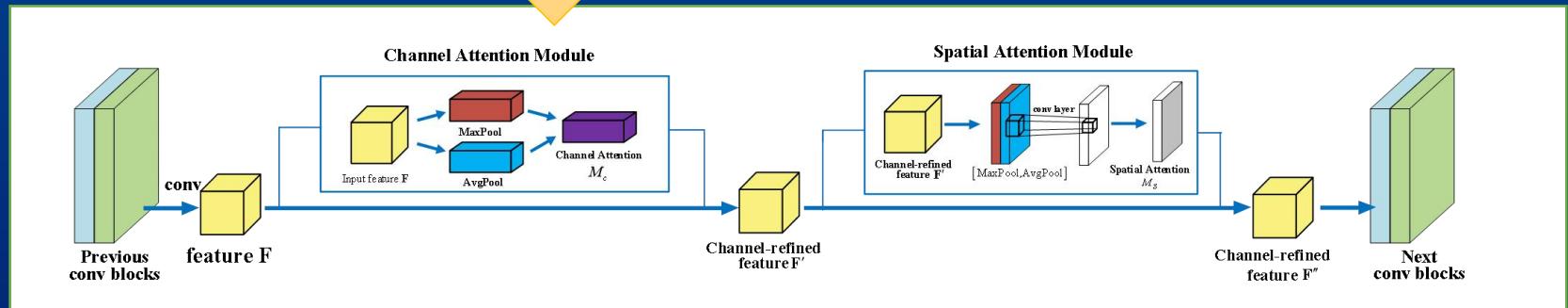
## • 04 创新设计 —— 首创三阶段生成对抗网络CLGN



➤ 多阶段网络：构建多阶段（粗略-局部-全局）生成网络CLGN，可从不同大小的感受野获取特征信息。

➤ 由深到浅：添加GAM全局注意力模块，将网络整体性能提高10.9%。

➤ 多次补全：保证语义和纹理一致性，修复后图像结构相似度高达86%。



**GAM全局注意力模块：**将信息量保持在最小并放大全局交互来提高网络性能

## • 05 实验结果 ——视觉效果对比



图1：使用不规则掩模（20-30%遮挡）人脸图片修复效果

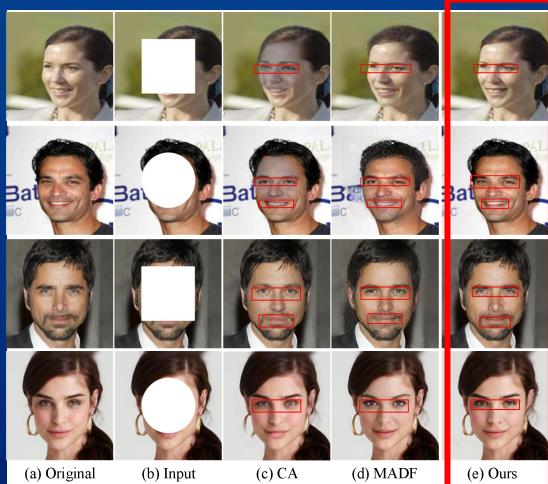
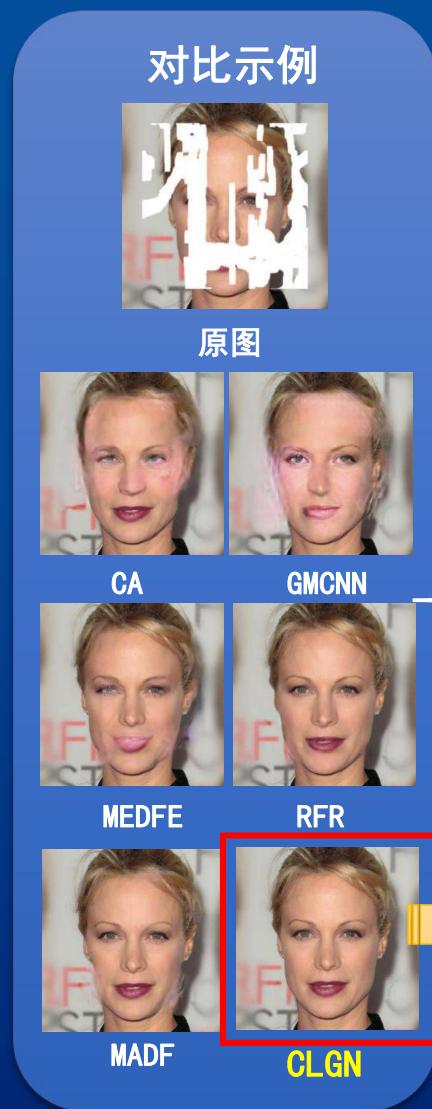


图2：使用规则掩模（矩形和圆形）覆盖的人脸图片修复效果



- 随机选取CelebA、CelebA-HQ数据集中共**200,000**个彩色人脸图像
- 160000**张用于训练，**40000**张用于测试
- 训练**200\*10 epoch**
- 全程耗时**1150**小时

### 纹理修复方法CA、MEDFE：

- 修复后图像存在灰色重影
- 人脸宽度有形变，错误率23.6%

### 生成式图像修复方法RFR：

- 出现部分区域模糊和伪影
- 修复区域边缘存在色差问题

### CLGN图像修复方法：

- 修复面部五官特征趋于完整；
- 孔洞修复纹理相似度高达**91.8%**
- 细节修复精度平均提高**18%**

# ● 06 实验结果 ——定量指标对比

表1：CLGN与其他六种方法规则掩膜修复指标比较

	遮损率	CA	GMCNN	MEDFE	RFR	LG-Net	Ours
损失函数 L1 (%) †	1-10%	1.77	1.54	1.43	1.57	0.44	<b>0.39</b>
	20-30%	5.28	3.01	3.72	3.74	2.45	<b>2.19</b>
	40-50%	7.92	<b>4.63</b>	7.64	6.51	5.31	<b>5.11</b>
峰值信噪比 PSNR ‡	平均	4.99	3.06	4.26	3.94	2.73	<b>2.56</b>
	1-10%	33.12	36.29	36.21	37.26	40.72	<b>42.25</b>
	20-30%	24.07	28.33	27.85	29.14	30.67	<b>31.66</b>
结构相似度 SSIM ‡	40-50%	21.11	26.08	23.5	25.23	26.09	<b>26.56</b>
	平均	26.1	30.23	29.19	30.54	32.49	<b>33.49</b>
	1-10%	0.971	0.977	0.99	0.99	0.995	<b>0.996</b>
图像相似度 FID †	20-30%	0.901	0.928	0.945	0.952	0.962	<b>0.986</b>
	40-50%	0.853	0.895	0.844	0.899	0.911	<b>0.913</b>
	平均	0.908	0.933	0.926	0.947	0.956	<b>0.965</b>
图像相似度 FID †	1-10%	2.14	0.82	0.79	0.85	<b>0.40</b>	0.41
	20-30%	6.82	2.26	3.21	2.73	2.11	<b>2.07</b>
	40-50%	12.39	<b>4.51</b>	7.19	5.22	4.6	<b>5.02</b>
图像相似度 FID †	平均	7.11	2.53	3.73	2.93	2.37	<b>2.5</b>

表2：CLGN与其他六种方法不规则掩膜修复指标比较

	遮损率	CA	GMCNN	MEDFE	RFR	LG-Net	Ours
损失函数 L1 (%) †	1-10%	1.86	1.14	1.02	1.59	0.46	<b>0.39</b>
	20-30%	5.33	3.05	3.68	3.58	2.38	<b>2.11</b>
	40-50%	7.84	<b>4.51</b>	7.65	6.44	5.27	<b>5.03</b>
峰值信噪比 PSNR ‡	平均	5.01	2.9	4.12	3.87	2.7	<b>2.51</b>
	1-10%	32.66	35.96	36.13	36.39	40.04	<b>41.53</b>
	20-30%	23.94	28.52	27.75	29.07	30.54	<b>31.33</b>
结构相似度 SSIM ‡	40-50%	21.98	25.89	23.47	25.09	26.01	<b>26.55</b>
	平均	26.19	30.12	29.12	30.18	32.19	<b>33.14</b>
	1-10%	0.971	0.984	0.99	0.991	0.995	<b>0.997</b>
结构相似度 SSIM ‡	20-30%	0.903	0.933	0.943	0.957	0.968	<b>0.987</b>
	40-50%	0.853	0.897	0.865	0.902	0.917	<b>0.921</b>
	平均	0.909	0.938	0.932	0.95	0.96	<b>0.968</b>
图像相似度 FID †	1-10%	2.06	0.85	0.84	0.86	0.39	<b>0.37</b>
	20-30%	6.97	2.24	3.17	2.67	<b>2.08</b>	2.11
	40-50%	12.42	4.56	7.12	5.21	4.61	<b>4.47</b>
图像相似度 FID †	平均	7.15	2.55	3.17	2.91	2.36	<b>2.31</b>

## 对比效果

损失函数L1 Loss

↓ 21.2%

图像平滑性max

峰值信噪比

↑ 13.5%

背景噪声min

结构相似度

↑ 16.7%

结构还原度高

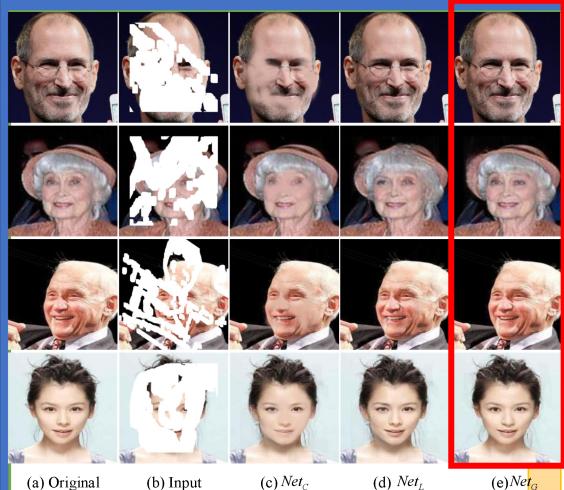
图像相似度

↓ 57.4%

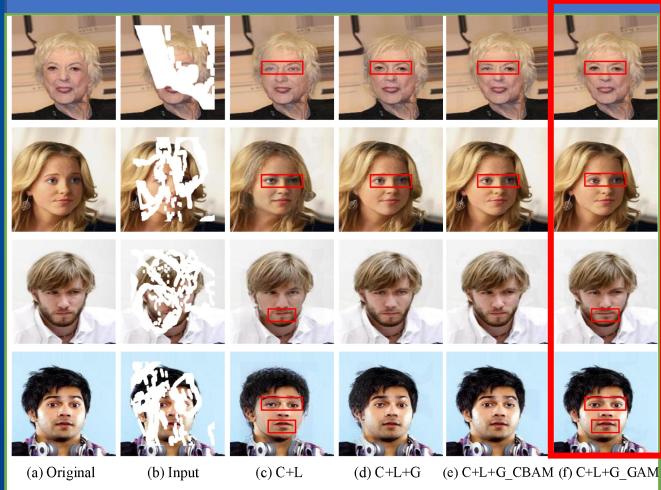
图像失真小

## • 07 实验结果 —— 消融实验

逐层网络性能对比



注意力模块性能对比



损失函数影响对比

损失函数	峰值信噪比	结构相似度	图像相似度	感知度量
重建损失	23. 01	0. 958	7. 62	0. 128
对抗损失	22. 93	0. 931	7. 27	0. 134
总变分损失	23. 04	0. 943	7. 31	0. 151
感知损失	22. 73	0. 907	7. 18	0. 133
风格损失	22. 98	0. 972	7. 23	0. 136
<b>共用</b>	<b>23. 14</b>	<b>0. 980</b>	<b>7. 11</b>	<b>0. 124</b>

三层网络叠加时结构相似度平均提升0. 36

添加注意力模块后峰值信噪比提升6. 23, L1范数误差减少2. 15

CLGN网络中使用的五种**损失函数**, 对提升网络性能有效, 消融实验测试结果对比各指标平均提高了1% - 6% (PSNR、SSIM、FID、LPIPS)



返璞归真  
FAN PU GUI ZHEN

## • 08 实验结果 ——问卷调查



- 在谷歌平台，收集有效问卷 2880份。历时3个月。
- 问卷有5个小题，分为单选题和打分题，限时3分钟。
- 旨在调研出6种传统图像修复方法与CLGN方法在人眼视觉下优劣性的对比。

图1&2:不同图像修复方法的支持率占比

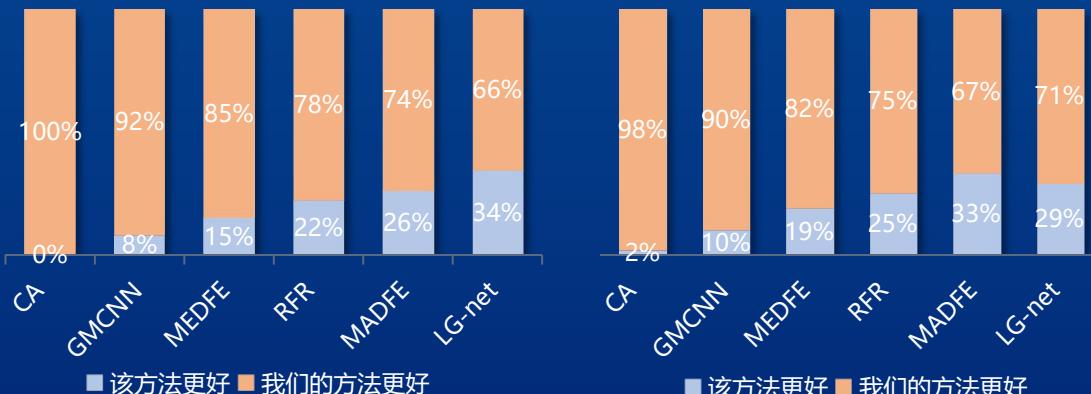
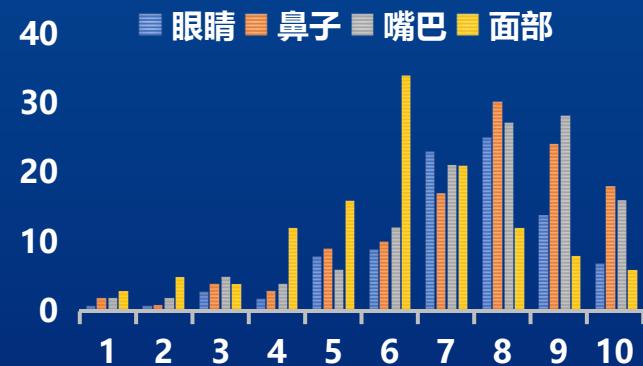


图3:五官修复情况评分柱状图



• 结论：平均有88.5%的用户认为相比于其他6种传统方法，CLGN方法修复出来的图片视觉效果更佳。

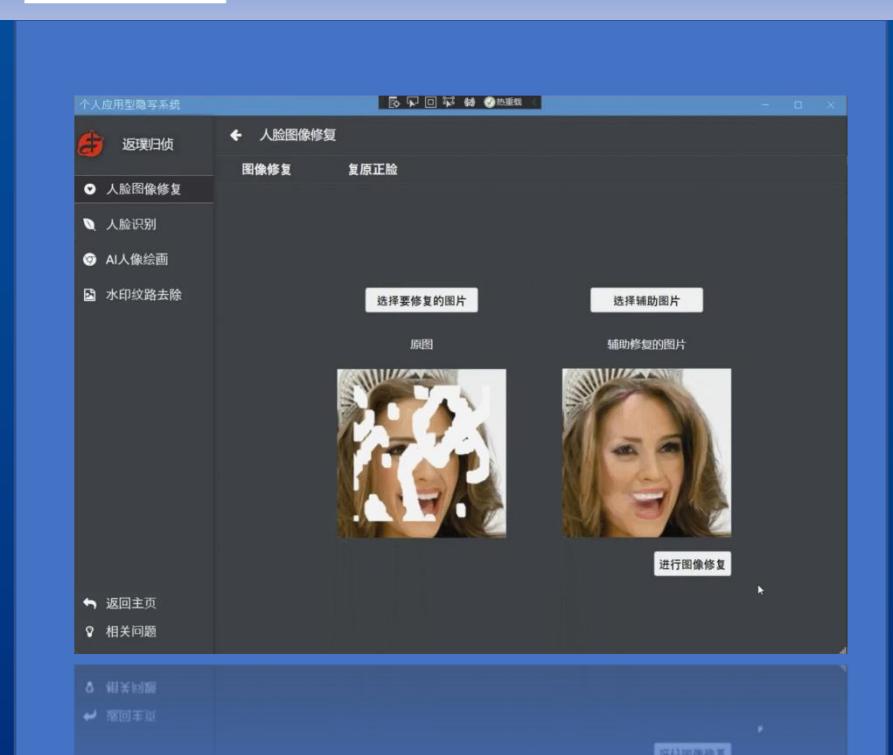


返璞归真  
FAN PU GUI ZHEN

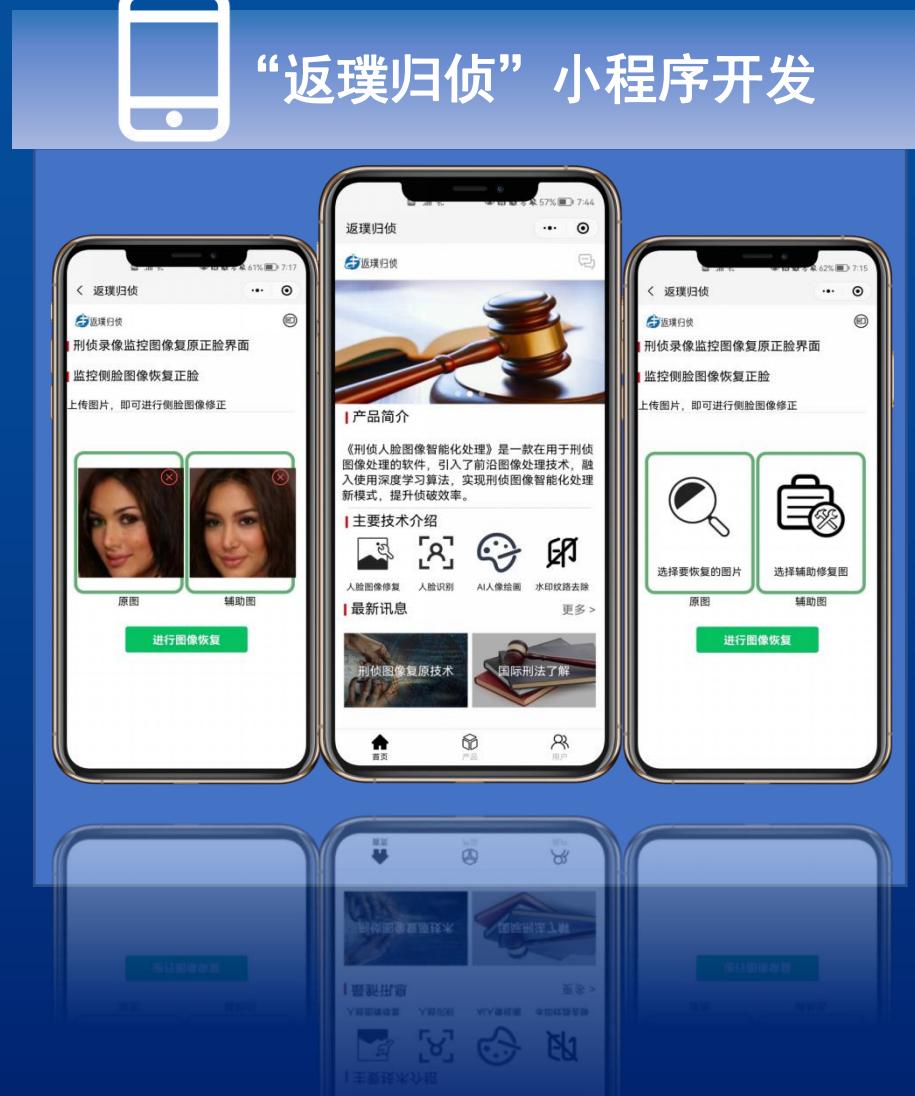
# ● 09 应用设计



## “返璞归真” PC端设计

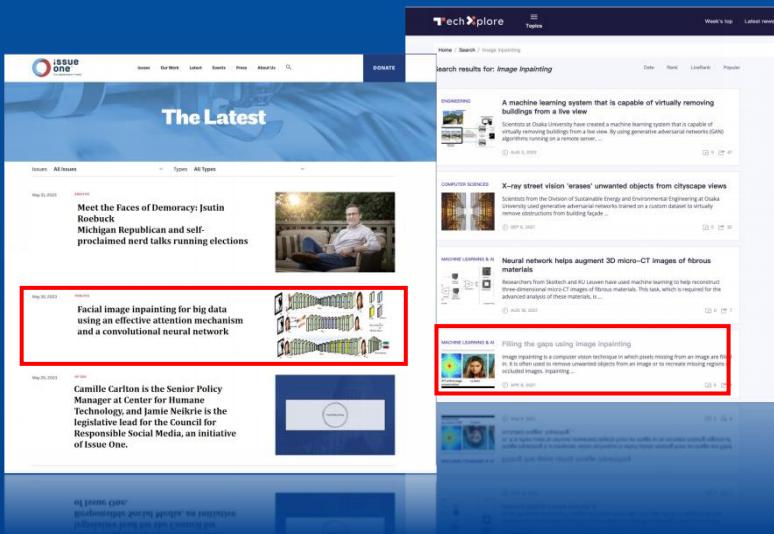


## “返璞归真” 小程序开发



# ● 10 业内认可

## 社会影响力——社会及媒体报道



## 学术影响力——计算机、信息领域权威专家认可

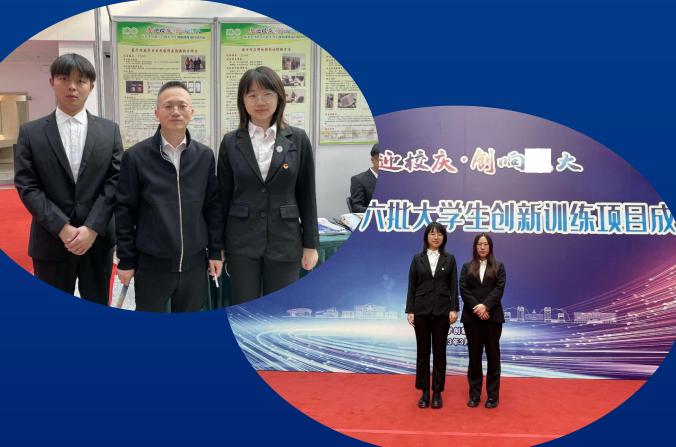


计算机软件国家工程  
研究中心副主任覃文军

“作品以人脸图像为研究对象，选题具有一定的前沿性，研究成果适用于刑侦领域，如人脸遮盖物去除、人脸去皱纹等，具有较强的应用价值。”

对作品的意义、技术水平、适用范围及推广前景作出评价	该作品以人脸图像为研究对象，选题具有一定前沿性和创新性，对国内外相关领域发展现状的归纳准确到位，在论文中建立了基于注意力机制和生成对抗网络的人脸修复方法，其研究成果适用于刑侦领域，如人脸遮盖物去除、监控侧脸图像恢复、人脸识别去皱等，具有较强的实用价值，同意推荐。
推荐者签字	

## 教育影响力——项目展览交流



“兴辽英才计划”  
科技领军人才刘金海



“该论文研究方法翔实，具有良好创新性；在应用方面，可以为公安、司法系统提供影像处理与分析技术，推广前景可期。”

对作品的意义、技术水平、适用范围及推广前景作出评价	该作品探讨了基于注意力机制和生成对抗网络的人脸图像修复方法，论文研究方法较为翔实，具有较好的创新性；在实际应用方面，该作品可以为公安、司法等系统提供影像处理与分析技术，如人脸图像缺失复原，生成不同年龄段对比图，模糊人脸清晰化等，推广前景可期。总体而言，该作品是一篇优秀的学术论文，有实际应用价值，同意推荐。
推荐者签字	

# 11 知识产权

SCI论文发表

1篇

EI论文发表

2篇

发明专利公开

2项

软件著作权

6项

第一作者，第一发明人  
均为组内本科生



## ● 12 所获成果 ——业内合作及获奖经历



### 业内合作



### 项目合作

- 君宏兆业（辽宁）文化有限公司 —— 人脸图像数据集/标签共享
- 杭州慧舍网络科技有限公司 —— CLGN隐写算法
- 杭州市政房地产有限公司 —— 提供“返璞归真”小程序

### 项目应用成果

- “该算法在刑侦人脸修复系统中，修复后的人脸失真率在 0.3% 以下，并能够在监控侧脸图像还原正脸、模糊变形人脸图像修复矫正等方面提供一定的技术支持。”

——杭州慧舍网络科技有限公司应用证明

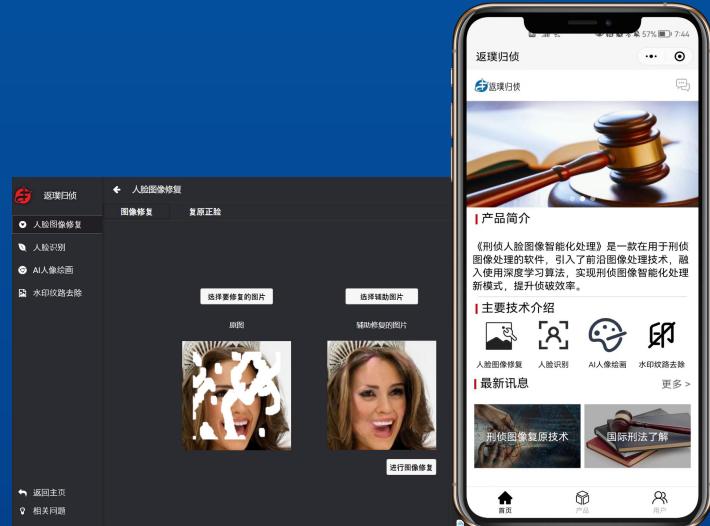
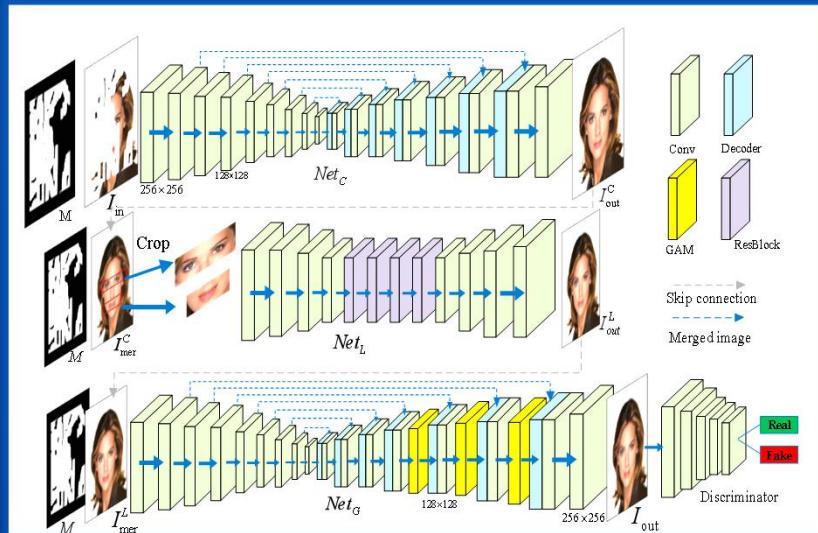


### 获奖经历



项目获2022“大学生创新创业训练计划”  
国家级优秀结题、最佳创意项目

## ● 13 总结



本论文提出的基于**生成对抗网络和注意力机制**的人脸图像修复方法，因其修复后**图像精度高、结构连贯、五官特征完整**等特点，可被应用于“返璞归侦”等人像处理工具中，提升人脸图像在刑侦、公安中的可用性。



# 人脸图像智能化处理研究

请各位评委老师批评指正！