用于跨分辨率人脸识别的属性指导的耦合GAN

Veeru Talreja \*，Fariborz Taherkhani \*，Matthew C.Valenti 和Nasser M.Nasrabadi

西弗吉尼亚大学Morgantown，西弗吉尼亚州，美国

vtalreja @ mix.wvu.edu，fariborztaherkhani @ gmail.com，valenti @ ieee.org，nasser.nasrabadi @ mail.wvu.edu

# 摘要

*在本文中，我们提出了一种新颖的属性指导的交叉分辨率（从低分辨率到高分辨率）人脸识别框架，该框架利用耦合的生成对抗网络（GAN）结构和对抗训练来发现低分辨率和高分辨率之间的隐藏关系。潜在的公共嵌入子空间中的高分辨率图像。*

*耦合的GAN框架由两个子网组成，一个子网专用于低分辨率域，另一个子网专用于高分辨率域。每个子网的目的是找到一个投影，该投影可以最大化公共嵌入子空间中两个特征域之间的成对相关性。除了将图像投影到公共子空间中之外，耦合网络还可以预测人脸属性以改善交叉分辨率人脸识别。特别，我们提出的耦合框架通过在训练过程中通过隐式匹配低分辨率和高分辨率图像的面部属性来利用面部属性来进一步最大化成对相关性，从而导致更具判别性的嵌入子空间，从而提高了交叉分辨率面部识别的性能。使用LFWA，Celeb-A，SCFace和UCCS数据集证明了我们的方法与最新技术相比的有效性。*

# 简介

~~面部生物识别技术可用于各种现代识别和监视应用程序，从银行和超市中的独立摄像头应用程序到执法应用程序甚至基于云的身份验证应用程序中的多台联网闭路电视。[40–45]。监视摄像机与对象之间的较大距离会导致捕获图像中的低分辨率（LR）面部区域。通常，在LR图像中人脸的判别性能会下降，这会导致为高分辨率（HR）图像开发的传统人脸识别算法的准确性大大下降。一个有效的面部识别算法即使对于LR面部也应表现良好，而不会显着降低识别精度。~~

arXiv:1908.01790v1 [cs.CV] 5 Aug 2019

~~与HR脸部图像相比，LR脸部具有自己独特的视觉特性。尽管LR面部图像中缺少许多视觉特征，但人类仍然能够注意到给定对象的LR和HR面部图像之间的相似性。这意味着如果人脑熟悉该对象的高分辨率图像或给定的身份，人脑的神经系统就能够恢复LR面部缺失的视觉特性[5]。受这一事实的启发，已经引入了几种LR人脸识别模型，这些模型通常可以分为两类：幻觉类和嵌入类。幻觉类别中的模型在识别之前从LR面孔重构HR面孔[3，11，18，19，35，46，51]。超分辨率的幻觉类别也用于其他应用[4]。例如，Kolouri和Rohde [18]提出了一种基于最优传输的单帧超分辨率方法，以自动构建HR面部表情的非线性拉格朗日模型。此后，通过探索完全适合给定LR数据的模型参数来改善LR面部图像。~~

~~基于幻觉的方法通常在识别重建的HR脸部图像方面取得可喜的结果。但是，~~幻觉模型中的超分辨率操作通常需要大量额外的计算，这通常会导致识别速度降低。与基于幻觉的方法相反，基于嵌入的方法通过利用各种外部面部上下文从LR面部提取特征。~~任等人。[29]引入了耦合内核嵌入，将具有不同分辨率的人脸图像隐式映射到无限空间中。然后，在此新空间中执行识别任务，以最小化分别由低分辨率空间和高分辨率空间中的其内核Gram矩阵获得的差异。凭直觉~~ 嵌入方法的主要步骤是将知识从HR面部图像转移到LR面部图像。但是，在这些方法中，必须小心谨慎，仅转移所需的知识，而不是将所有知识从HR域转移到LR域。

除了在高分辨率图像和低分辨率图像之间共享知识外，诸如面部属性之类的软生物特征也可以用作补充信息，以改善跨分辨率的面部识别模型。面部属性先前已在不同的面部识别应用程序中与面部生物特征一起使用[2]。

在本文中，我们提出了一种基于属性指导的深度耦合学习框架的跨分辨率人脸识别嵌入模型，该模型使用生成对抗网络（GAN）在共同的嵌入子空间中来寻找潜在的高分辨率和低分辨率图像特征之间的隐藏关系。该框架还利用卷积神经网络（CNN）权重共享，然后使用特有的权重来学习每个特定面部属性的代表特征。具体来说，我们的耦合框架利用面部属性进一步最大化了低分辨率域和高分辨率域之间的相关性，这导致了更具区分性的嵌入子空间，从而增强了跨分辨率人脸识别任务的性能。另外，在我们的方法中我们还预测了多任务范例中的跨分辨率面部识别的低分辨率图像的属性。多任务学习试图通过利用两个任务之间的知识共享来同时解决相关任务[37-39]。总而言之，我们的主要贡献是：

* 使用耦合GAN和多个损失函数的新型属性导向的交叉分辨率（低分辨率到高分辨率）人脸识别模型。
* 多任务学习框架，可预测低分辨率面部图像的面部属性。
* 使用四个不同的数据集进行了广泛的实验，并将所提出的方法与最新方法进行了比较。

# 相关工作

低分辨率人脸识别有两种方法。基于幻觉的方法[3，19，35,46,51]在人脸识别之前重建高分辨率人脸，而基于嵌入的方法则通过使用嵌入技术直接从低分辨率和高分辨率的人脸中提取潜在特征。杨等。[51]利用稀疏表示来同时执行识别和幻觉，以合成特定于人的低分辨率面部表情版本，而不会明显降低识别率。在[46]中，提出了一种算法，该算法通过稀疏表示来识别带有特定字典的人脸，该字典包括许多自然和面部图像。此外，诸如[3]和[19]中提出的模型之类的深层模型可以从低分辨率面部生成强烈逼真的高分辨率图像。不过，

而不是重建高分辨率的面部，一种更直接的方法是将低分辨率的面部嵌入到不同的外部环境中，以在分辨率降低时检索丢失的信息[5]。受此启发，人们提出了嵌入方法，将高分辨率和低分辨率的人脸都转换为用于匹配的集成特征域[5、8、12、20、23、24、33、47、48、50、52]。在[23]中，多维缩放用于学习通用变换矩阵，以共同变换低分辨率和高分辨率训练面部图像的面部特征。另一方面，王等。[48]通过深度学习方法解决了非常低分辨率的识别问题。在[9]中，CNN与基于流形的轨迹比较技术一起用于视频中的低分辨率人脸识别。

# 生成对抗网络

生成对抗网络（GANs）已广泛用于不同的计算机视觉应用中。GAN由两个竞争网络组成，分别是生成器G和鉴别器D。GAN的目标是训练生成器G从训练噪声分布*p z*（*z*）生成样本，以使鉴别器D无法将合成样本与实际数据区分开*y*，分布*p 数据*。发电机*G ^*（*ż* ; *θ 克*） 是一个微函数它映射的噪声变量*ž*使用参数的数据空间*θ 克*。在另一方面，鉴别器*d*（*。* ; *θ d*）也是一个可微函数，它试图使用实数据之间的二元分类区分*Y*和*G ^*（*Ž*）。具体来说，生成器和鉴别器在两人minimax游戏中相互竞争，以使Jenson-Shannon差异最小化[6]。GAN 的损失函数*L*（*D，G*）给出为：

*大号*（*d，G*）= *è y*〜*P 数据*（*Y*） [日志*d*（*ÿ*）]

（1）

+ *ë ž*〜*P Ž*（*ż*） [日志（1 - *d*（*g ^*（*ż*）））]

GAN的目标（两个玩家的minimax游戏）由下式给出：

极大极小*大号*（*d，G*）=极大极小[ *ë y*〜*P 数据*（*Y*） [日志*d*（*ÿ*）]

*ģ d ģ d*

+ *ë ž*〜*P Ž*（*ż*） [日志（1 - *d*（*g ^*（*ż*）））]]

（2）

|  |
| --- |
| 图1：建议框架的框图。 |

条件GAN是GAN的另一种变体，其中生成器和鉴别器都以附加变量*x*为条件。该附加变量可以是任何种类的辅助信息，例如离散标签[22]，文本[28]或图像[10]。条件GAN的损失函数为：

*大号Ç*（*d，G*）= *è y*〜*P 数据*（*Y*） [日志*d*（*Y* | *X*）]

（3）+ *è Ž*〜*P Ž*（*ż*） [日志（1 - *d*（*g ^*（*Ž* | *X*）））] *。*

条件GAN的目标是与（2）中相同的两人minimax游戏，损失函数为*L c*（*D，G*）。此后，我们将条件GAN的目标表示为*O cGAN*（*D，G，y，x*），由下式给出：

*ø cGAN*（*d，G，Y，X*）=极大极小[ *ë y*〜*P 数据*（*Y*） [日志*d*（*Y* | *X*）]

*g ^ d*

+ *ë ž*〜*P Ž*（*ż*） [日志（1 - *d*（*g ^*（*Ž* | *X*）））]] *。*

（4）

# 提出的方法

在本节中，我们描述了提出的用于跨分辨率人脸识别的方法。

幻觉方法：生成器将每个低分辨率图像上采样到高分辨率。

嵌入方法：使用生成模型将高分辨率和低分辨率图像投影到一个共同的低维嵌入子空间。受到GANs [6]成功的启发，我们在多任务学习中探索对抗网络，以将低分辨率和高分辨率图像投影到公共子空间进行识别/预测面部属性。

如图1所示，所提出的方法包括由两个子网络组成的耦合框架，其中每个子网都是由生成器和鉴别器组成的GAN架构。生成器通过对比损失函数耦合在一起。每个生成器还在多任务学习中负责预测面部属性。除了对抗性损失和对比性损失外，我们还建议使用基于VGG 16架构的感知损失以及*L* 2重构误差来指导子网络。这是因为优化中的感知损失有助于实现逼真的图像重建[13]。

## 深度耦合框架

我们的方法的目的是进行低分辨率人脸识别。低分辨率图像和高分辨率图像的匹配在公共嵌入子空间中执行。因此，我们使用包含两个子网的耦合框架：低分辨率（LR）网络和高分辨率（HR）网络。

LR网络由GAN（生成器+鉴别器），属性预测器和VGG-16感知网络组成，

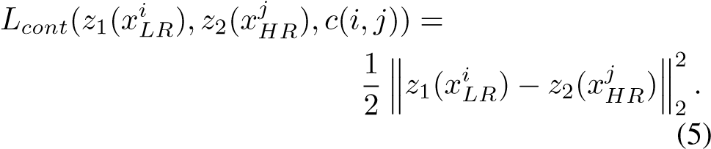
而HR网络由GAN（生成器+鉴别器）和属性预测器组成。

对于生成器，我们使用了U-Net网络[30]来更好地捕获低级特征并克服由于深度网络而导致的梯度消失问题。受[10]的启发，我们使用了PatchGAN。基于补丁的鉴别器可确保保留高频细节（当仅使用*L* 1损耗时，这些高频细节通常会丢失）。我们提出的方法的最终目标是在一个公共的嵌入子空间中找到全局深层潜在特征，这些特征代表了低分辨率及其对应的高分辨率面部图像之间的关系。为了找到两个域之间的这个公共子空间，我们通过对比损失函数*L cont* [1] 耦合两个生成器。

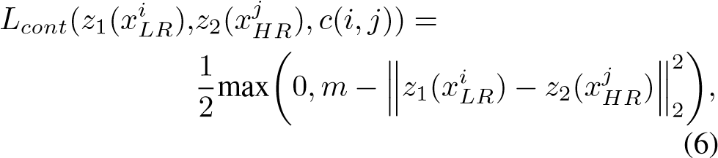
将该损失函数（*L cont*） 最小化，以便在一个公共嵌入子空间中将真正的对（即具有自己的对应HR图像的LR图像）彼此相互靠近，同时让非匹配对彼此远离（增大类间距离，减小类内距离）。

损失函数定义和上一篇文章中的对比损失一样，表示输入的LR脸部图像，并且*x j HR*表示输入的HR图像。*Ç*（*I，J*）是一个二进制标记， 如果和*X Ĵ HR*属于同一类，它等于0，不属于同一类则等于1。令*z* 1（*。*）和*z* 2（*。*）表示生成网络中提取嵌入的函数，分别输入的HR和LR图片转换为公共子空间。

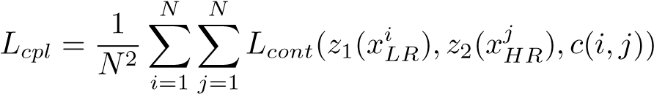
如果*c*（*i，j*）= 0 （即真实对），则对比损失函数（*L cont*）给出为：



类似地，如果*c*（*i，j*）= 1 （即冒名顶替者对），那么对比损失函数 （*L cont*） 为：



其中*m*是对比边距，用于“收紧”约束。因此，用于耦合子网的总损耗函数由*L cpl*表示，并表示为：

*，*

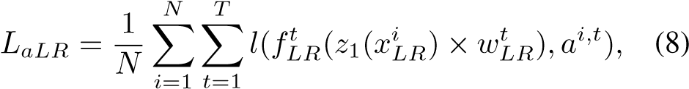
（7）

其中N是训练样本数。使用耦合损失的主要动机是因为它隐式使用类标签，因此能够找到有区分性的嵌入子空间，而其他某些度量标准（例如欧几里得距离）则可能不是这种情况。此判别性嵌入子空间对于LR图像与HR图像的匹配以及属性预测任务很有用。

## 属性预测任务

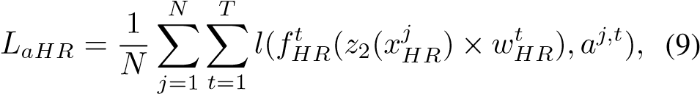
除了跨分辨率的面部识别，我们提出的方法的另一个重要目标是使用LR或HR面部图像预测属性。但是，通过单独学习多个CNN实现这两个任务并不是最佳选择，因为不同的目标可能具有共同的特征并且具有隐藏的关系，可以利用这些关系联合优化目标。联合优化的概念已在[53]中使用，他们训练CNN用于人脸识别，并利用特征进行属性预测。因此，对于此任务，我们使用从公共嵌入子空间得到的各自的特征集（*z* 1（*x i LR*）或*z* 1（*x j HR*）），也可以预测给定图像的属性。同样，我们的网络在不同的属性预测任务之间共享其大部分参数，以增强多任务范式中识别任务的性能。

对于属性预测任务，将LR或HR图像作为输入提供给网络，以预测一组属性。考虑所述输入是由*X iLR*的LR图像表示，其中该图像的类标签由下式给出*l* *i* ∈ L for（i = 1 *，* ... *，N ）*，其中*Ñ*是训练样本的数目。让*Ť*为不同的面部的属性的数目，a*I，t*表示用于训练样本i的ground truth属性标签。在这种情况下，使用来自公共嵌入子空间的特征集，属性预测损失函数为：



其中 *f LR t* （*。*）是在LR生成器的瓶颈层上对于属性*t*的二进制分类器，如图1所示。该分类器是通过使用损失函数*l（例如交叉熵）*和*w LRt*代表分类器的权重参数来学习，并且针对每个面部属性任务分别学习这些参数。

同样，我们可以考虑其他子网（HR网络），并使用HR图像执行相同的过程。 对于给定的HR图像，HR网络还使用来自公共嵌入子空间的特征来预测一组HR属性的面部集合。因此，用于HR图像的面部属性预测的损失函数为：



其中的符号类似于（8），但对应于HR网络。总的属性预测损失为：

*L a* = *L aLR* + *L aHR 。* （10）

## 生成对抗损失

令*G LR*和*G HR*表示分别从输入LR和HR图像合成相应的LR和HR图像的生成器。令*D LR*和*D HR*分别表示LR和HR GAN的鉴别器。我们已经利用GAN损失函数[6]来训练生成器和相应的鉴别器，以确保鉴别器无法将生成器的合成图像与相应的ground truth图像区分开。此外，从图1可以看出，生成器*G LR*和*G HR*尝试使用分别以输入LR和HR图像为条件的网络来生成LR和HR图像。耦合GAN的总损失为：

*L GAN* = *L LR* + *L HR ，* （11）

其中 *L LR* 和*L HR*分别表示LR和HR网络的GAN损失函数，分别表示为：

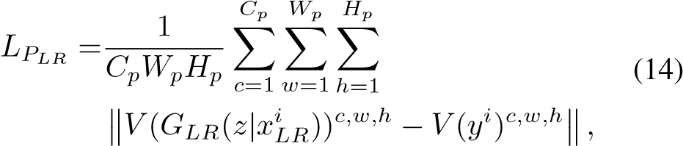
 （12）

*L HR* = *O cGAN*（*D HR ，G HR ，y j ，x jHR*）*，* （13）

其中，函数*O cGAN*由（4）给出。*x i LR*（*x j HR*）是输入LR（HR）图像，而*y i*（*y j*）表示实际LR（HR）数据。注意，实际LR（HR）数据*y i*（*y j*）和*x i LR*（*x j HR*）给出的网络条件是相同的。

## 感知损失

在[13]中引入了感知损失函数，用于风格转换和超分辨率。在[13]中，使用仅从预训练卷积神经网络中提取的高级图像特征表示之间的误差来学习网络参数，而不是仅依赖于*L* 1或*L* 2重构误差。同样，在我们提出的方法中，使用预先训练的VGG-16 [34]网络提取高级特征（ReLU3-3层）将感知损失添加到LR网络中，以及真实和合成特征之间的*L* 1距离用于引导发生器*G LR其中*LR网络的特征的感知损失为：



其中 *y i* 是ground truth LR图像， 是LR生成器的输出。*V*（*。*）代表VGG-16网络的特定层，其中层的维度由 *C p*，*W p*和*H p*给出。我们仅将感知损失应用于LR网络，以生成有助于识别的更shaper的LR图像。

同样，我们将感知损失也用于属性预测，以测量合成图像和真实图像的面部属性之间的差异。我们为LR和HR网络的属性应用了感知损失。为了从给定的HR图像中提取属性，我们在12个带注释的面部属性上微调了预训练的VGG-Face [27]，这些属性在表2中显示。此后，我们利用该属性预测器来测量LR和HR网络的属性感知损失：

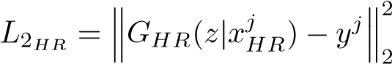
其中A（。）是微调的VGG-Face属性预测器网络。总的属性感知损失是LR网络（*L pa LR*）和HR网络（*L pa HR*）的感知属性损失之和：

*大号PA* = *大号PA LR* + *大号PA HR 。* （17）

## *L* 2重建损失

*L* 2重建损失根据合成图像和相应的真实图像之间的欧几里得距离来度量重建误差，并针对LR和HR网络定义如下：

 （18）

 *。* （19）

总的*L* 2重建损失函数由下式给出：

*L* 2 ＝*L* 2 *LR* + *L* 2 *HR 。* （20）

## 总体目标函数

提出的方法中用于学习网络参数的总体目标函数为所有上述定义的损耗函数的总和：

*大号TOT* = *大号CPL* + *λ* 1和*大号一个* + *λ* 2 *大号GAN*

（21）

+ *λ* 3 *大号P LR* + *λ* 4 *大号PA* + *λ* 5 *大号* 2 *，*

其中*L cpl*是耦合损失函数，*L a*是总属性预测损失函数，*L GAN*是总生成对抗损失函数，*L P LR*是LR网络的感知损失，*L pa*是总感知属性损失函数和*大号*2是总的重建误差。*λ* 1 *，λ* 2 *，λ* 3 *，λ* 4 *，λ* 5是可调节的超参数权衡不同损耗项。

# 实验与结果

在本节中，我们通过对四个数据集进行各种实验来证明该方法的有效性：Wild-a中的带标签的面孔（LFWA）[49]，CelebFaces属性数据集（CelebA）[21]，监视摄像机的面孔（SCFace） [7]和不受约束的大学生（UCCS）数据集[32]。我们已将我们提出的方法与不同数据集上的六个最新方法进行了比较：VLRR [48]，DCA [8]，LRRFW [20]，D-align [24]，SHSR [35]和SKD [5 ]。此外，我们进行了消融研究，以证明我们网络中每个损失功能的有效性。

## 数据集

CelebA由202,599张图像组成，分别经过训练，验证和测试分割，分别约有162,000、20,000和20,000张图像。整个数据集对应于大约10,000个身份（每个身份20张图像），没有身份重叠。图像带有40个面部属性，例如“波浪发”，“胖”，“秃头”，“男性”等注释。但是，我们提出的方法仅使用12个属性（如表2所示）。我们使用数据集的预裁剪版本，其中面部图像使用手工标记的关键点对齐。常规HR分辨率下的图像大小等于178×212 。我们将图像下采样为88×108 ，68×84 ，48×58 。

LFWA共有13232张图像，包含5749个身份，具有预定义的训练和测试划分，将整个数据集划分为大约两个相等的分区。每个图像都用CelebA数据集中的40个相同属性进行注释。图像被归一化到224×224 为HR图像和向下取样，以96×96 ，64×64 ，32×32 。

SCface数据集由130个对象组成，每个对象具有一个HR正面人脸图像和多个HR图像，并使用不同质量的监视摄像机从三个距离（分别为4.2m，2.6m和1.0m）捕获。为了与以前的方法公平比较[8]，随机选择50个受试者进行训练，其余80个受试者进行测试。如[8]，我们固定在HR图像128×128 和下采样到64×64 ，32×32 和16×16 为LR图像。

UCCS数据集是在不受约束的条件下获取的极具挑战性的数据集。按照[48]中的实验设置，我们对180个对象的子集执行评估，其中每个对象具有25个或更多图像。我们总共获得5,220张图像，并使用4,200张图像进行训练，其余1,020张图像用于测试。为了公平比较，我们将裁剪后的脸部区域归一化为80×80 作为HR，然后对16×16的LR图像进行降采样。对于未使用属性注释的数据集SCFace和UCCS，我们使用最新的混合目标优化网络（MOON）[31]来生成基本事实属性。

## **训练细节**

**如前所述，我们已经将U-Net网络实现为用于LR和HR网络的生成器和基于补丁的判别器。**整个架构已在Pytorch中实现。收敛便于网络收敛，所有的超参数都设置为1，除了*λ* 3，和*λ* 4，被设置为0.5。

对于Adam优化器[17]，我们使用的批量大小为6，一阶动量为0.5，学习率0.0004​​。我们对生成器使用了ReLU激活函数，对鉴别器使用了0.25斜率的Leaky ReLU。为了微调属性预测器网络VGG-Face用于属性感知损失，我们选择12个属性（如表2所示，来自LFWA数据集）。

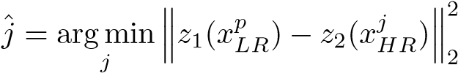
（21）中的复杂损失使得难以直接训练整个网络。为了解决这个问题，我们采用了阶段性学习策略，将训练数据中的信息逐渐呈现给网络。具体来说，我们首先通过不同时更新其他任务来贪婪地优化每个任务。在对每个任务进行初始化之后，我们通过共同优化所有任务对整个网络进行微调。

对于训练，我们需要真对和假对。真对和假对是使用相同/不同subject的LR和HR图像构建的。我们通过使用相同数量的真实对和伪造对来平衡训练集。

## 测试提出的方法

我们提出的方法的主要目标是使用来自公共嵌入子空间的相应特征集，将测试LR图像与HR图像库进行匹配。

在测试期间，给定的probe LR图像*x p LR*通过LR网络，生成嵌入*z* 1（*x p LR*）。类似地，来自gallery的HR图像通过HR网络传递，执行相同的过程，并且针对每个gallery图像*x j HR* 生成*z* 2（*x j HR*）。最终，通过为计算*z* 1（*x p LR*）和所有gallery HR图像的*z* 2（*x j HR*） 之间的最小欧几里得距离来执行人脸识别：

 *。* （22）

因此，是来自gallery的HR图像，给定的待匹配LR图像是*x p LR*。将正确分类的probe数与probe总数之比作为识别率。

另外，通过将特征集*z* 1（*x p LR*）通过LR网络的属性预测器，LR网络还可用于给定LR probe图像的面部属性预测。预测的面部属性可用于缩小在大型HR图像库中进行识别的搜索范围。

**40.97**

**5**

**32.7**

**2**

**29.3**

**22.35**

**18**

**.3**

**16.48**

**Top-1**

**Top-5**

**0**

**5**

**10**

**15**

**20**

**25**

**30**

**35**

**40**

**45**

**Error Rate**

|  |
| --- |
| 图2：针对不同数据集的低分辨率图像的rank-n的识别精度的CMC曲线。 |

图3：使用UCCS数据集的VLLR（蓝色），SKD（绿色）和我们的方法（黄色）的Top-1和Top-5错误率比较。

**100**

**95**

**Recognition Accuracy**

**90**

**85 96  64 32**

### 低分辨率图像尺寸

图4：对于不同尺寸的LR图像，使用LFWA数据集的SKD（蓝色）和我们的方法（黄色）的识别精度（％）比较。

## 性能 评估

我们对提出的方法进行了评估，并在四个不同数据集上与其他最新方法进行了比较。图2使用LFWA，CelebA和SCFace数据集提供了我们提出的方法从rank 1到rank5对不同分辨率图像的识别精度。我们可以清楚地看到，所提出的方法在LFWA和CelebA数据集上有非常好的性能。然而，SCFace具有更有挑战性面变化

表1：评价等级-1识别精度（％上SCFace）。

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 模型 | 1区 | Dist-2 | Dist-3 | 平均 |
|  |  |  |  |  |
| SKD | 43.50 | 48.00 | 53.50 | 48.33 |
| 我们的方法 | 44 *。*81±0 *。*36 49 *。*60±0 *。*41 54 *。*30±0 *。*23 49 *。*57±0 *。*39 | | | |

1

0

0.2

0.4

0.6

0.8

False Positive Rate (FPR)

0

0.2

0.4

0.6

0.8

1

True Positive Rate (TPR)

Total loss

L

cpl

+

L

2

+

L

GAN

+

L

P

LR

L

+

pa

L

cpl

+

L

2

图5：与消融研究相对应的ROC曲线。

，而CelebA和SCFace图像是在典型的商业监视环境中拍摄的，这导致了与LFWA和CelebA相比，SCFace数据集的识别性能较低，如图2所示。

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 表2：使用CelebA数据集的属性预测准确性（％）比较。   |  |  |  |  |  |  |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | |  | 双下巴 | 胖乎乎的 | 眼镜 | 男 | 皮肤苍白 | 胡子 | 嘴巴稍微张开 | 年轻 | 面带微笑 | 山羊皮 | 秃 | 金发 | | 人力资源输入（净额A） | 90.8 | 90.1 | 96.5 | 97.5 | 89.2 | 92.8 | 90.1 | 85.3 | 91.3 | 93.8 | 97.4 | 92.7 | | LR输入（网络A） | 51.3 | 50.1 | 56.5 | 55.3 | 46.4 | 53.1 | 51.6 | 49.2 | 55.4 | 56.9 | 62.8 | 59.7 | | LR输入（微调网络A） | 71.6 | 69.3 | 74.6 | 72.9 | 68.6 | 71.4 | 70.3 | 69.5 | 75.8 | 76.0 | 88.5 | 81.4 | | LR输入（建议方法） | 83.2 | 82.6 | 89.3 | 91 | 83.3 | 88.1 | 86.6 | 78.9 | 87.2 | 88.9 | 93.6 | 88.4 | |

表1列出了使用距相机的不同距离（**对应于不同分辨率的图像**）的SCFace数据集与不同最新方法对Rank-1识别率的比较。我们可以看到，我们提出的方法比最新交叉分辨率人脸识别嵌入式方法模型SKD [5]优越1 *。*31％，1 *。*60％，0 *。*80％和1 *。*Dist-1，Dist-2，Dist-3和平均值分别为22％。我们还可以观察到，我们的模型甚至比最先进的幻觉模型SHSR [35]高出约33％这三个距离的平均值。我们还使用UCCS数据集将建议的方法与VLRR [48]和SKD [5]进行了比较。Top-1和Top-5的错误率比较如图3所示。UCCS还是一个极富挑战性的数据集，在完全不受约束的条件下捕获了面部图片。由于这个原因，该数据集的错误率非常高。但是，我们提出的方法比其他两种方法具有更好的性能，因为它的错误率至少为3 *。*4％和1 *。*对于Top-1和Top-5认可度为8％。从图4中我们还可以注意到，即使对于不同分辨率的LFWA数据集，我们提出的方法也优于SKD。

从性能评估中，我们观察到我们提出的具有对比损失函数并利用面部属性将不同域（LR和HR）转换为通用判别性嵌入子空间的耦合框架优于其他嵌入技术（如SKD和D-Align）。它还显示了利用多种损失函数进行交叉分辨率人脸识别的功效。损耗功能的相对重要性已在消融研究中详细介绍（第5.6节）。

## 低分辨率的属性预测

该方法的优点之一是可以用于LR面部图像的属性预测。为了说明我们提出的方法对LR脸部图像的属性预测的有效性，我们针对4种不同的场景执行了属性预测：

1、使用基于VGG-Face的属性预测器对HR图像进行属性预测，在Sec 4.4中表示为Net A 。4.4。

2、使用属性预测器Net A对LR图像进行属性预测。

3、在这种情况下，我们首先通过这将被称为“微调网络A”。

4）在最后一种情况下，我们从LR网络测试for属性预测，以进行LR测试图像的属性预测（见图1）。我们已经使用以下方法对Celeb-A数据集68×88 作为LR图像进行了此实验。

表2中列出了使用Celeb-A对上述4个场景的12个属性的属性预测结果。从表2中可以看出，我们的方法在预测两个数据集的LR图像的属性方面显示出最佳性能。用LR图像对Net A进行微调有助于改善其性能，但是它的性能不如我们的方法。此外，我们的LR网络属性预测器的性能与HR图像的Net A性能相当。

## 消融研究

（21）中定义的目标函数包含多个损失函数：耦合损失（*L cpl*），属性预测损失（*L a*），感知损失（*L P LR*，*L pa*），*L* 2重建损失（*L* 2）和GAN损失（*L GAN*）。在本节中，我们研究了不同损失函数的相对重要性以及在我们提出的方法中使用它们的好处。对于本实验，我们使用提出的方法的不同变体，并使用LFWA数据集（64×64 LR图像）进行评估。变体是：1）使用耦合框架进行交叉分辨率人脸验证，仅耦合损失和*L* 2重建损失（*L cpl*+ *L* 2）; 2）使用具有耦合损失，*L* 2重建损失，GAN损失和知觉损失（*L cpl* + *L* 2 + *L GAN* + *L P LR* + *L pa*）的耦合框架的交叉分辨率面部验证；3）使用我们所有损失函数（*L cpl* + *L* 2 + *L GAN* + *L P LR* + *L pa* + *L a*）的框架进行交叉分辨率人脸验证。

我们使用我们框架的上述三个变体，并使用公共嵌入子空间中的特征绘制接收器工作特性（ROC）曲线，以用于交叉分辨率人脸验证任务。从图5可以看出，生成对抗损失和知觉损失（红色曲线）有助于提高交叉分辨率验证性能，而添加属性预测损失（蓝色曲线）有助于进一步提高性能。进行此改进的原因是，将面部属性丢失与对比丢失一起使用会导致更具判别性的嵌入子空间，从而带来更好的面部识别性能。这也表明，属性预测和面部识别的多任务学习是有用的，并且有助于跨分辨率的面部识别任务。

# 结论

我们提出了一种新颖的框架，该框架采用耦合GAN并利用面部属性进行交叉分辨率面部识别。耦合的GAN包括两个子网，这些子网将低分辨率和高分辨率图像投影到一个公共的嵌入子空间中，其中每个子网的目标是在投影过程中最大化低分辨率和高分辨率图像之间的成对相关性。此外，我们在训练过程中通过隐式匹配低分辨率和高分辨率图像的面部属性来利用面部属性来进一步最大化成对相关性。我们在四个标准数据集上全面评估了我们的模型，结果表明我们的模型在交叉分辨率人脸识别方面明显优于其他最新模型。另外，

1. 原文

提供更好的翻译建议