TWO-BRANCH的深度卷积神经网络架构来实现低分辨率人脸识别

Erfan Zangeneh，Mohammad Rahmati和Yalda Mohsenzadeh

*抽象*-我们提出了一种使用深度卷积神经网络（DCNN）的低分辨率人脸识别的新颖耦合映射方法。所提出的体系结构由两个分支的DCNN组成，以通过非线性变换将高分辨率和低分辨率的人脸图像映射到公共空间。对应于高分辨率图像变换的分支由14层组成，将低分辨率面部图像映射到公共空间的另一个分支包括连接到14层网络的5层超分辨率网络。反向传播相应的高分辨率和低分辨率图像的特征之间的距离以训练网络。我们对提出的方法进行了FERET数据集的评估，并与最新的竞争方法进行了比较。我们广泛的实验评估表明，该方法显着提高了识别性能，尤其是对于分辨率非常低的探针人脸图像（识别精度提高了11.4％）。此外，它可以从其相应的低分辨率探针图像重建高分辨率图像，就视觉质量而言，该图像可与最新的超分辨率方法相媲美。

*索引词* -低分辨率人脸识别，超分辨率方法，耦合映射方法，深度卷积神经网络

# 一，引言

在这里，我们专注于解决在提供高质量图像库时识别低分辨率探头面部图像的问题。有三种解决此问题的标准方法。1）将图库图像降采样到探测图像的分辨率，然后执行识别。但是，这种方法不是最佳的，因为会丢失高分辨率画廊图像中可用的其他区分信息。2）第二种方法是从低分辨率图像中获得更高分辨率的探针图像，然后将其用于识别。这些超分辨率技术中的大多数旨在根据视觉质量重建高质量的图像，并且并未针对识别性能进行优化[9]。此类别中一些众所周知的方法是[10] – [13] 3）最后，第三种方法将LR探测器图像和HR画廊图像同时转换为一个公共空间，其中对应的LR和HR图像距离最近。[14] – [17]是这种方法的众所周知的方法。图1总结了解决低分辨率人脸识别（LR FR）问题的三种通用方法。

arXiv:1706.06247v1 [cs.CV] 20 Jun 2017

在本文中，我们使用第三种方法，并提出了一种使用深度卷积神经网络（DCNN）的方法来找到低分辨率和高分辨率对人脸图像之间的公共空间。尽管先前的工作使用线性方程作为目标函数来找到两个投影矩阵，但我们的工作还是发现了从LR和HR到公共空间的非线性变换。在我们提出的方法中，在公共空间中变换后的低分辨率和高分辨率图像的距离被用作目标函数来训练我们的深度卷积神经网络。我们提出的方法还可以重建良好的HR人脸图像，这对于识别任务来说是最佳的。我们在FERET数据库上评估了所提出方法的有效性[18]。我们的结果表明，与其他现有技术相比，该方法在低分辨率人脸识别中的匹配性能显着提高，并且对于非常低分辨率的探针图像，改进效果更加明显。这项研究的主要贡献可以概括为：

* 我们提出了一种新颖的非线性耦合映射结构，该结构使用两个深层卷积神经网络将低分辨率和高分辨率面部图像投影到一个公共空间中。
* 与其他最先进的竞争方法相比，我们提出的方法具有更高的识别精度，尤其是在探针图像的分辨率极低的情况下。
* 由于其体系结构中嵌入了超分辨率CNN，因此我们提出的耦合映射方法还提供了低分辨率输入图像的高分辨率版本。
* 我们的方法相比，我们需要的空间要少得多

到使用深度卷积神经网络（例如VGGnet）的典型面部识别方法[19]。此功能使其适用于内存较低的常规系统。

接下来，我们首先回顾第2节中在超高分辨率方法和耦合映射方法领域在低分辨率人脸识别领域中提出的先前工作。在第3节中，我们介绍我们提出的方法，网络体系结构和训练过程。最后，我们在第4节中介绍了实验评估结果，并在第5节中给出了讨论和结论。

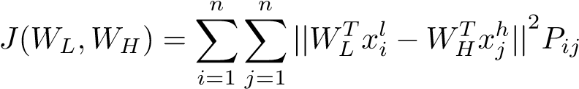
# 二。以前的工作

在本节中，我们简要回顾了低分辨率人脸识别文献中的相关工作，并介绍了深度卷积神经网络。

低分辨率人脸识别：为了解决探针图像和画廊图像之间的不匹配，大多数研究都集中在超分辨率方法上。这些方法的目的是从LR输入获得HR图像，然后将获得的HR图像用于识别。一些超分辨率研究建议使用面部先验图像重建。Chakrabarti等人提出的学习方法。[20]的超分辨率使用内核主成分分析来得出有关面部类别的先验知识。为了获得良好的重建效果，刘等。[11]提出了一种用于从LR输入图像中幻化出HR面部图像的两步统计建模方法。贝克[21]还提出了一种基于面部先验的面部幻觉方法。Freeman等。文献[22]将分段马尔可夫网络训练为超分辨率预测模型。杨等。[13]使用压缩感测从低分辨率输入图像中重建超分辨率图像。Zou等。[10]提出了一种超分辨率方法，该方法将低分辨率的人脸图像聚类，然后将与分配的聚类相对应的投影矩阵将LR图像映射到HR空间。他们提出了两个单独的投影矩阵，以实现最佳的可视化和识别目的。在[13]中，作者提出了一种稀疏编码方法来根据其相邻的图像斑块找到LR输入斑块的表示。然后使用相同的表示系数来重建目标 [10]提出了一种超分辨率方法，该方法将低分辨率的人脸图像聚类，然后将与分配的聚类相对应的投影矩阵将LR图像映射到HR空间。他们提出了两个单独的投影矩阵，以实现最佳的可视化和识别目的。在[13]中，作者提出了一种稀疏编码方法来根据其相邻的图像斑块找到LR输入斑块的表示。然后使用相同的表示系数来重建目标 [10]提出了一种超分辨率方法，该方法将低分辨率的人脸图像聚类，然后将与分配的聚类相对应的投影矩阵将LR图像映射到HR空间。他们提出了两个单独的投影矩阵，以实现最佳的可视化和识别目的。在[13]中，作者提出了一种稀疏编码方法来根据其相邻的图像斑块找到LR输入斑块的表示。然后使用相同的表示系数来重建目标

基于相应的相邻HR补丁的HR补丁。

关于LR FR的另一类作品称为耦合映射方法。这些方法使用由相同对象的HR图像和LR图像组成的训练集学习变换。由于训练数据，我们的目标是寻找一种最小化转化LR和HR特征向量之间的距离，转换和分别。大多数耦合映射方法使用线性目标函数，如下所示[23]： 

 （1）

其中， *n* 是训练图像的数量，{ *x h i* } *i n* = 1 ，分别是HR和LR图像的相应提取特征。 *W L*和*W H*表示线性

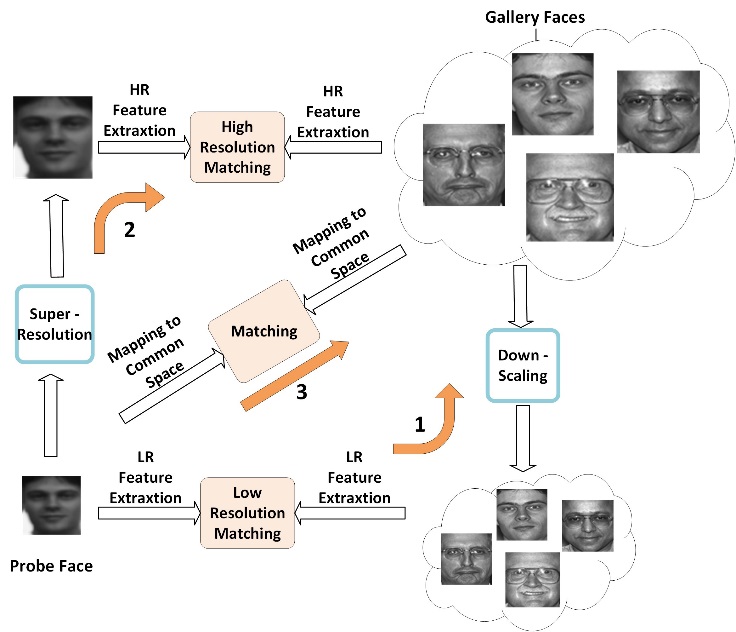
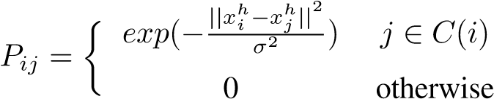
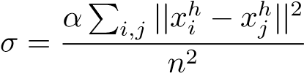


图1.低分辨率人脸识别的三种通用方法。

低分辨率和高分辨率特征向量分别映射到公共空间。*P*是*n* × *n*惩罚加权矩阵，它保留原始特征空间中数据点之间的局部关系，并且在数据点的邻域上定义如下：

 （2）

此处， *C*（*i*）包含高分辨率空间中*x h i*的*k个*最近邻居的索引，并且*σ*是高斯函数宽度，其定义为

 （3）

|  |
| --- |
| 图2.我们的方法概述。HR和LR图像，分别和的m，n表示尺寸*M> N*。 |

其中，*α*是一个规模parameter.Since假设HR功能空间具有更多的判别信息，目标的目标函数上面找到类似HR特征空间的共同特征空间。最后，在优化了上述目标函数之后，*W L*和*W H*将发现低分辨率和高分辨率图像可以分别通过这些映射转换为公共空间。体育亨宁斯等。[14]提出了一个联合目标函数，旨在优化超分辨率和人脸识别。尽管与两步法相比，该方法提高了识别精度，但其优化过程却很慢。主要是因为必须针对每个注册针对每个测试图像执行其优化过程。在[15]中，作者使用了多分辨率人脸图像的奇异值分解（SVD）将低分辨率图像映射到高分辨率空间。此外，该方法提高了幻觉和识别精度。黄等。[24]提出了一种方法，该方法为低分辨率探针和高分辨率画廊图像找到一个公共空间，并提供一个目标函数来保证在新的公共空间中的可分辨性。Biswas等。[16]使用多维缩放变换学习来找到低分辨率和高分辨率投影矩阵。优化问题的目标函数在公共空间中强制使一类低分辨率和高分辨率图像对之间的距离与该类高分辨率图像对的距离相同。黄等。[24]使用规范相关分析（CCA）将低分辨率和高分辨率图像投影到一个公共空间，在该公共空间中，低分辨率图像及其对应的高分辨率图像尽可能接近。后来，任等人。[25]采用耦合的非线性核将LR和HR面部图像对映射到一个无限的公共子空间中。在Zhou et al。中也提出了一种耦合线性映射方法。[17]使用经典判别分析。Shi等。[26]提出了一个优化目标函数，该函数包括与LR / HR一致性，类内紧凑性和类间可分离性相关的三个术语。张等。[27]介绍了耦合边际判别映射（CMDM）方法。该方法利用来自HR图像的成对均值样本对之间的高斯相似度来构造类内相似度矩阵。类间相似度矩阵由同一类HR图像的样本对之间的高斯相似度定义。此外，张等。[27]解决了耦合映射问题作为本征分解问题，当面部被遮挡时有助于获得良好的识别性能。Mudunuri等。[28]提出了一种耦合映射方法，该方法首先通过检测眼睛来对齐人脸，然后计算探针人脸的SIFT描述符以将其变换为公共空间。然后，使用立体匹配成本函数来保留变换后的空间中不同照明，姿势和分辨率之间的距离。

总之，耦合映射方法比超分辨率方法具有更好的识别性能，但是这些方法的目的并不是要从低分辨率输入图像中重建高分辨率图像。另一方面，超分辨率方法的主要目的是为了可视化目的重建高质量图像，而不一定提供更好的识别精度。

深度卷积神经网络：尽管卷积神经网络（CNN）于三十年前首次提出[29]，但自2012年引入AlexNet [30]以来，深层CNN便爆炸性地流行，尤其是由于其在计算机视觉领域的成功。促成这种流行的主要因素是[31]：

* 参加顶级大学的最佳实验室应对计算机视觉挑战，例如ILSVRC [32]和PASCAL VOC [33]
* 轻松访问较大尺寸的数据，例如ImageNet

[34]。

* 引入了更有效的激活函数，例如整流线性单元（ReLU）[35]和指数线性单元（ELU）[36]，它们有助于DCNN更快地收敛。
* 存在诸如NVIDIA TITAN black X的现代GPU，以及高效的深度学习框架，例如Caffe [37]和Tensorflow [38]。

在下一部分中，我们提出一种使用深度卷积神经网络的非线性非线性映射到公共空间的耦合映射方法。我们的方法类似于使用深度卷积神经网络的其他成功方法，受益于上述优点。除了提供高识别性能之外，所提出的方法还从低分辨率输入图像产生高分辨率图像。

# 三，建议的方法

由于难以解决非线性优化问题，因此采用线性变换对以前的耦合映射方法（如第2节所述）中的目标函数进行建模。但是，将低分辨率和高分辨率非线性转换为公共空间可能会导致更好的性能。在这里，我们提出了一种非线性耦合映射方法，该方法使用两个深层卷积神经网络（DCNN）从低分辨率探针图像和高分辨率画廊图像中提取特征并将它们投影到公共空间中。我们使用基于梯度的优化来最小化公共空间中映射的HR和LR图像对之间的距离

|  |
| --- |
| 图3.我们提出的方法的两个分支中的两个深度卷积神经网络的体系结构。 |

通过误差的反向传播来更新DCNN的权重。

图2显示了我们提出的体系结构的概述。在训练阶段，我们使用的训练图像集包含成对的同一个人的低分辨率和高分辨率图像，这些图像在不同的光照，姿势和表情条件下可能会有所不同（不一定是仅具有不同分辨率的同一幅图像） 。在下一节中，我们将详细介绍所提出的方法的体系结构。

## A.网络架构

我们的方法有一个two-branch的结构，其中一个将高分辨率图像投影到公共空间，另一个将低分辨率图像映射到该公共空间。在我们的方法中，我们使用称为VGGnet的DCNN[9]。该网络最广为人知的配置有十六层，其中有十三层卷积层和三层完全连接的层。VGGnet的最后一个完全连接层用于特定分类任务。在我们方法的顶部分支（图2）中，我们删除了该VGGnet的最后两个完全连接的层，并将其称为特征提取卷积神经网络（FECNN）。我们方法顶部分支的输入图像是高分辨率图像（*I i h*），必须为224×224 尺寸（无论输入图像大小与224×224 不同，我们都使用传统的双三次插值方法来获得所需大小）。最后一层的输出是具有4096个元素的特征向量。

在我们方法的底部分支中，我们使用先前用于超分辨率低分辨率图像的DCNN，之后是第二个网络，该网络的结构与顶部分支中的网络相似。第一个子网的结构与Dong等人提出的DCNN类似。[31]，但是我们将这种架构从三层扩展到了五层，尽管作者表明，在重建图像的视觉质量方面，三层架构和五层之间没有区别，我们发现层数从三层增加到5可以提高我们方法的识别性能。我们将底部分支超分辨率网络（SRnet）称为第一个子网。第一个子网的输出被馈送到第二个子网（FECNN）。因此，），必须使用传统的插值方法将输入LR图片其插值为224×224 的大小。此外，SR子网的输出是大小为224×224 的图像。

如上所述，除了最后两个完全连接的层之外，FECNN网络具有与VGGnet相同的体系结构。尽管超分辨率和特征提取卷积神经网络（SRFECNN）具有18个卷积层和一个完全连接的层，但SRFECNN中使用的权重总数远少于VGGnet。表一显示了SRFECNN的所有已使用权重。即使我们提议的SRFECNN包括18个卷积层，但由于与VGGnet相比，全连接层的数量较少，因此其权重数量比VGGnet（141 *M*重量）。因此，在测试阶段，当我们需要将SRFECNN权重加载到内存中时，我们提出的方法需要的空间比VGGnet少得多。这是一个重要的功能，使我们提出的方法适用于内存较低的系统。

## B.共同子空间学习

我们分三个阶段对网络进行了培训，概述如下：

* 首先，我们使用在人脸数据集[19]上经过训练的VGGnet，然后删除最后两个完全连接的层，因为它们特定于网络在其上进行训练的分类任务。我们称此网络为预训练的FECNN，并在我们架构的顶部和底部分支中使用了它。

表一

Ñ 棕土所用权重的层SRFECNN。

|  |  |  |
| --- | --- | --- |
| 图层集 | 参量 | 重量数 |
| 转换0 1 | *F* = 9×9  *深度* = 96 | 3×9×9×96 = 23328 |
| 转换0 2 | *F* = 1×1  *深度* = 64 | 96×1×1×64 = 6144 |
| 转换0 3 | *F* = 1×1  *深度* = 48 | 64×1×1×48 = 2928 |
| 转换0 | *F* = 1×1  *深度* = 32 | 48×1×1×32 = 1536年 |
| 转换0 | *F* = 5×5  *深度* = 3 | 32×5×5×3 = 2400 |
| 转换1（2转换） | *F* = 3×3  *深度* = 64 | 2（3×3×3×64）= 3456 |
| 转换2（2转换） | *F* = 3×3  *深度* = 128 | 2（64×3×3×128）= 147456 |
| 转换3（3转换）  转换4（3转换） | *F* = 3×3  *深度* = 512 | 3（256×3×3×512）=  3538944 |
| *F* = 3×3 3（512×3×3×512）=  转换5（3转换）  *深度* = 512 7077888 | | |

* 第二步，我们使用高分辨率和低分辨率人脸图像对的数据库训练了底部分支的SRnet。实验评估部分介绍了使用的数据集的详细信息。
* 第三步是主要培训阶段。我们合并了SRnet和FECNN这两个子网，并将包含低分辨率和高分辨率同一对人的成对训练数据库分别输入到底部和顶部分支。

我们将顶部分支FECNN网络和底部分支SRFECNN视为将高分辨率图像和低分辨率图像投影到4096维公共空间的两个非线性函数：

|  |  |
| --- | --- |
| *φ 喜* = *˚F ħ*（*我IH*） | （4） |
|  | （5） |

其中*我IH* ∈ *[R 中号* × *中号*和*予IL* ∈ *[R Ñ* × *Ñ*该 *Ñ <M*。

在训练的这一阶段，*F H*（*I i h*）被认为是固定的并且没有变化，但是*F L*（*I i l*） 被训练为最小化同一空间中同一对象的低分辨率图像和高分辨率图像之间的距离。为此，该距离作为误差反向传播到了底部分支网络（FECNN和SRnet）中。

对于所有训练图像对，都重复了主要训练过程多次。我们降低了所有层的学习率，以微调在前两个训练阶段中获得的权重。但是，FECNN的第一层的学习速率小于它的最后一层，因为在特定问题中，DCNN的最后一层具有关于该问题的更多判别信息，并且它的第一层具有可以稀疏改变的更一般的特征[39]。

## C.重建输入图像

另外，我们的方法可以从低分辨率探针图像重建高分辨率图像。底部分支的第一个子网用于超分辨率，以从低分辨率探头表面生成高分辨率面部图像并馈入FECNN。在测试阶段，将低分辨率探针图像输入到底部网络后，我们可以从SRnet的最后一层提取相应的高分辨率面部图像。

## D.测试阶段

在测试阶段的第一阶段，所有高分辨率画廊图像都被馈送到顶部分支网并映射到公共空间，而探测图像则被馈送到底部分支网。

探针图像的标签由以下公式确定

*标签*（*I i l*）= *标签*（*I k h*） （6）

其中k由

*k* =精氨酸（7）

*Ĵ*

其中*我我升* 是低分辨率探头图像，*我ķ ħ* 是*ķ 吨 ħ*高分辨率库图像和*Ñ ģ*表示高分辨率脸部像册的图像的数目。

# IV。实验评估

人脸识别的两个主要任务是人脸识别和验证。在面部识别中，将查询的面部与galllery面部数据库进行比较以确定其身份。在人脸验证中，将验证所要求的查询人脸身份。在本节中，所有实验结果都属于人脸识别任务。实验旨在回答以下问题：

* 与最新的超分辨率方法和耦合映射方法相比，该方法的性能如何？
* 所提出的方法在不同分辨率的探针图像上执行的鲁棒性如何？
* 所提议的表情，光照和年龄变化的方法的鲁棒性如何？
* 当超分辨率子网从体系结构中排除时，建议的方法如何执行？
* • 所提出的方法从低分辨率探头重建高分辨率人脸图像的效果如何？
* • 如果将SR子网从其体系结构中排除，则在培训阶段中所提议方法的收敛性如何受到影响？

为了证明我们提出的方法的有效性，我们将我们的方法的面部识别性能与最新的竞争方法进行了比较，其中包括一种超分辨率（判别性超分辨率（DSR）方法[10]）和三种耦合映射方法（耦合局部保留映射（CLPM）[23]，相干特征上的非线性映射（NMCF）[24]和多维缩放（MDS）[16]方法））。

## A.数据说明

在本节中，我们描述了用于训练和评估我们提出的方法的数据集。

**训练数据集**：我们用于训练的数据集的详细信息在表II中列出。总共，我们使用了45315个面部图像，这些图像在姿势，表情，照度和年龄上都有变化。从FERET数据集[40]中，我们在训练中使用了10585张图像，在评估阶段使用了其余（3541张图像）。

**评估数据集**：我们在FERET [18]人脸数据库的一部分上进行了评估。FERET人脸数据集包含来自1199个人的14126人脸图像。分配此数据集的子集（包括3541个图像）进行评估。该数据集包括四个探针类别，每个类别都分配了一个图库集。所有画廊的脸部图像都是正面的。以下说明了四个探针类别的特征：

* 第一个探头类别称为*FB*，其中包括1195个正面面部图像。它的画廊集包含1196个表情各异的正面图像。
* 第二个探针类别（称为*重复* I）包含FERET数据库中的所有重复正面图像（722个图像）。图库与包含1196张图像的*FB*相同。
* 第三类称为*fc*，其中包括当天拍摄的194张图像，但相机和照明条件不同。图库与包含1196张图像的*FB*相同。
* 第四类称为*duplicateII，*由重复的探针图像组成，这些探针图像至少相差一年，并采集了相应的图库图像（不同的年龄条件）。为画廊*duplicateII*探针是库含有864个图像其他类别的子集。

## B.训练阶段

我们使用了预训练的VGGnet权重[19]并删除了最后两个完全连接的层来构建FECNN。同样，在训练我们的两个分支架构之前，我们在表II中描述的训练数据集上训练了SRnet。对于SRnet训练，我们首先从所有训练图像中对人脸进行降采样，以使数据集中对应的HR图像的LR人脸。SRnet包括五个卷积层，我们使用45315对LR和HR面部图像对网络进行了训练。分别训练FECNN和SRnet之后，我们连接了预先训练的SR和FECNN子网。然后，我们使用45315面训练了我们提出的体系结构。在训练阶段的主要部分中，我们降低了底部分支中每一层的学习率，以针对耦合映射目的在训练中微调底部网络。

## C.针对表达，光照和年龄变化的鲁棒性

在本实验中，我们在第IV-A节中描述的四类FERET评估数据集上评估了我们提出的方法。由于*FB*图像具有不同的表达条件，因此*fc*集包含具有不同表达条件的探针图像

表二

大号用于训练数据集的IST及其描述

图像数量的条款及其条件的变化

SUCH AS E：表达，我：照明，和病人：POSE 。 \* P 租赁注意

THAT FERET 数据集包含14126个映像和我们使用10585

IMAGE培训，其余的，3541个影像，用于评估。

|  |  |  |
| --- | --- | --- |
| 资料库 | 图片数量 | 强调 |
| 300瓦[41] | 600 | 在野外，  勘探与生产 |
| 海伦[42] | 2330 | 在野外，勘探与生产（E＆I＆P）的差异很大，并且具有遮挡作用 |
| IBUG [41] | 135 | 在野外，  勘探与生产 |
| 空军[43] | 250 | 在野外，  勘探与生产 |
| 佐治亚理工学院人脸数据库[44] | 750 | E＆I＆P的巨大差异 |
| LFW [45] | 13233 | 在野外，勘探与生产和规模的差异很大 |
| UMIST [46] | 564 | 灰度，以及姿势和种族的变化 |
| 耶鲁B [47] | 5760 | 灰度和P＆I的变化 |
| AT＆T [48] | 400 | 时间，眼镜和E＆I的变化 |
| 假装[40] | 14126 \* | 外观随时间变化以及P＆I＆E |
| CK + [49] | 10708 | 变动市盈率 |

照明条件和*重复II*集包含与相应画廊图像相比具有不同年龄条件的探测图像，我们也可以针对这些变化评估我们提出的方法的鲁棒性。在本实验中，尺寸为72×72 像素的HR脸部图像与两只眼睛的位置对齐。通过对对齐的HR脸部图像进行下采样和平滑操作，可以生成尺寸为12×12 像素的LR 图像。

图4显示了我们方法和四个竞争方法DSR [10]，MDS [16]，NMCF [24]和CLPM [23]的累积匹配曲线（CMC）。等级*k*的累积匹配分数是面部识别度量，其被定义为当HR画廊图像的*k个*最近邻居中的至少一个与LR探针图像属于同一个人时，探针图像的识别精度。图4中显示的结果表明，我们的方法的识别性能明显优于其他现有技术。图4.a描绘了*FB*上的累积匹配曲线数据集。如我们在第IV-A节中所述，此数据集仅在表达方面包括与图库图像不同的探针图像。我们提出的方法在第1级中的识别精度为91.8％，而竞争方法的最佳性能属于CLPM [23]，其识别精度为90.1％。我们提出的方法优于竞争方法，相差1.7％。图4.b描述了*fc*数据集上的CMC结果。与图库图像相比，此数据集中的探测图像的照明度有所不同。我们提出的方法在所有等级上都优于竞争方法。

|  |
| --- |
| 图4.我们提出的方法与CLPM [23]，MDS [16]，DSR [10]和NMCF [24]在识别率方面的比较。累积匹配曲线 |

（a）FB，（b）fc，（c）重复I和（d）重复II数据集。

很棒。与*fc*数据集上的最佳竞争方法相比，我们的方法在第1名中表现出5.6％的增长。这基本上显示了即使在不同光照条件下，深度卷积神经网络在特征提取和泛化方面的效率。尽管我们的方法的性能对光照的变化具有鲁棒性，但与*FB*相比，其他竞争方法的性能在*fc*数据集上显着下降。*DuplicateI*包含与图库条件类似的图像，但是表达略有差异。在该数据集上，我们方法的性能接近最佳竞争方法（DSR [10]）（图4.c）。该*duplicateII*与图库图像相比，包含的图像具有不同的年龄条件。我们提出的方法以5.2％的识别精度胜过排名1的最佳竞争方法（此处为MDS [16]）（图4.d）。同样，这表明我们提出的方法可以抵抗年龄变化。

综上所述，我们提出的方法在*FB*，*fc*和*重复II*探针图像上显示出最佳性能，而在*重复* I数据集上接近最佳性能。同样，我们的方法显示出针对表情，照度和年龄变化的鲁棒性，如图4.b和d所示。

表三

Ç OMPARISON RANK的1 识别精度在不同的PROBE图像分辨率。

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | 6×6 | 12×12 | 24×24 | 36×36 |
| 千次曝光出价[23] | 64.4％ | 90.1％ | 93.4％ | 95.2％ |
| MDS [16] | 57.3％ | 87.4％ | 90.2％ | 92.2％ |
| NMCF [24] | 60.3％ | 84.4％ | 88.4％ | 91.1％ |
| DSR [10] | 69.4％ | 88.5％ | 90％ | 93％ |
| 我们的方法 | 80.8％ | 91.8％ | 96.7％ | 98.8％ |

## D.对不同探头分辨率的评估

在这里，我们评估了我们提出的方法在分辨率非常低的探测图像上的有效性。在此实验中，我们在*FB*探针数据集上比较了我们的方法和最新方法的性能，该数据集的所有探针面都与图库面相似，但表达略有差异。因此适合研究分辨率变化的影响。我们考虑了四种不同的分辨率，6×6 ，12×12 ，24×24 ，和36×36 。每次我们训练SRnet

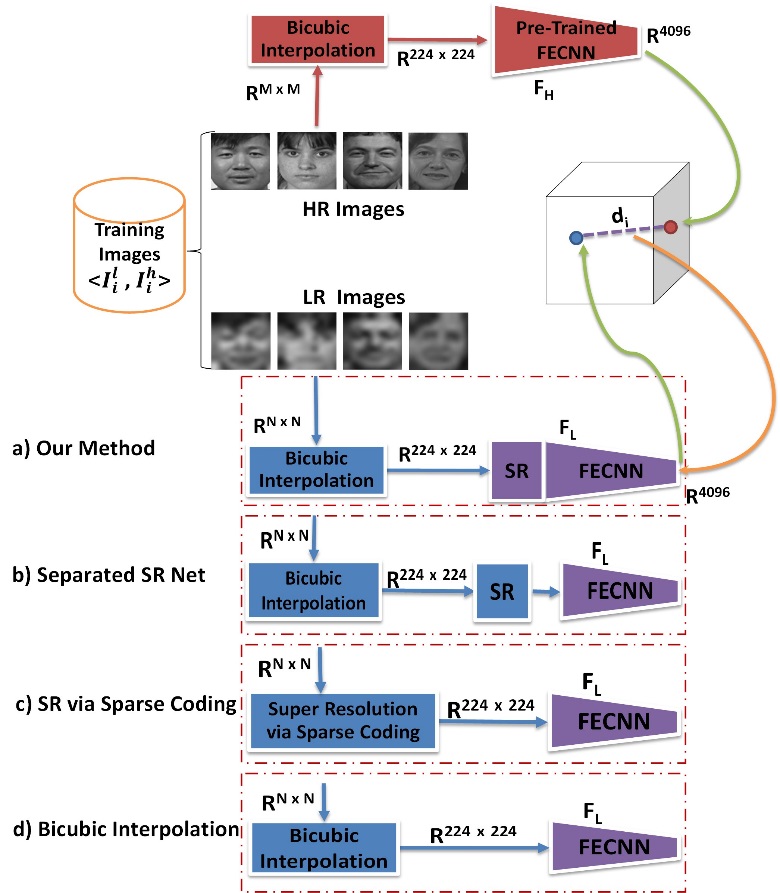


图5.具有不同超分辨率模块的配置 训练阶段涉及紫色模块。a）我们方法的配置。

b）SR子网与SRFECNN分开。c）将稀疏编码用于SR [13] d）仅使用双三次插值。

在具有降低的图像分辨率的训练数据上分别进行分离，然后将SRnet连接到FECNN，并分别在每种分辨率条件下对我们提出的方法的底部分支进行重新训练。表III显示了在*FB*集上评估的不同分辨率条件下，与竞争方法相比，我们方法的1级识别精度。可以看出，我们提出的方法在所有四个分辨率条件下都优于所有竞争方法。最显着的改进（11.4％）是在6×6的极低分辨率上，我们提出的方法优于DSR [10]，这是专门为识别极低的面部图像而提出的方法。

## E. SR子网的作用

如前所述，底部分支网络由两个网络SR和FECNN组成。在培训阶段，SR和FECNN网络都参与了主要培训阶段。在此实验中，我们旨在研究使用SRnet的影响及其微调对我们方法的识别性能的影响。图5显示了三种不同的配置，我们将我们提出的方法与它们进行了比较。我们提出的方法配置在图5.a中进行了描述，其中SR和FECNN子网都在主要训练阶段进行了训练。在图5.b所示的配置中，SRnet在底部分支中与FECNN分离，并且在主要训练阶段中，SRnet的权重保持固定。图5.c所示的配置采用稀疏编码[13]方法代替

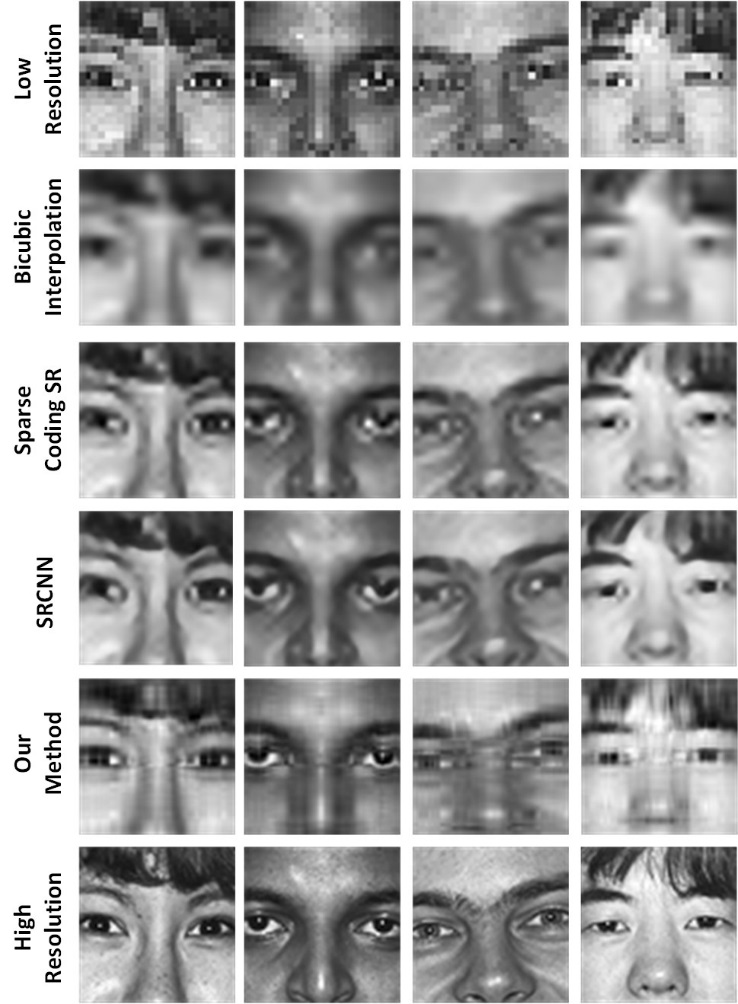


图6.通过图5中的不同配置重建的面孔

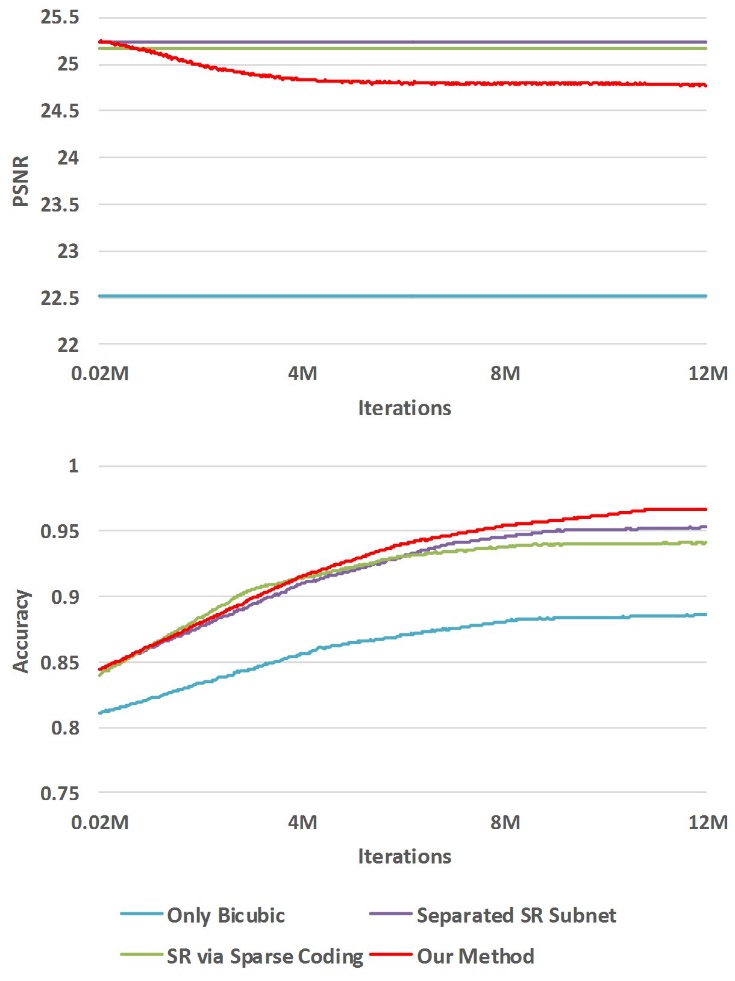
SRnet。同样，在主要训练阶段仅对FECNN进行训练。图5.d所示的配置仅使用双三次插值法将低分辨率输入图像映射到尺寸为224×224 的图像，因此不使用超分辨率网络。因此，在训练阶段，仅更新FECNN权重。表IV显示了四种不同配置的等级1识别精度（见图5）。这些结果说明，在配置中使用SRnet可以提高性能（请参阅表IV的第二行）。此外，将SRnet纳入主要培训阶段可以大大提高识别性能（我们在表IV中建议的方法）。特别是当探头组的分辨率很低时，识别

表四

ç OMPARISON RANK的1 识别的准确性不同的SR 模块配置在不同探头的图像分辨率。



|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 只有双三次的 | 66.7％ | 82.1％ | 88.6％ | 93.9％ |
| 分离的SR子网 | 75.4％ | 89.7％ | 95.3％ | 97.9％ |
| 通过稀疏编码进行SR | 73.9％ | 88.4％ | 94.1％ | 97％ |
| 我们的方法 | 80.8％ | 91.8％ | 96.7％ | 98.8％ |



|  |
| --- |
| 图7.比例因子为3时，在PSNR，SSIM和WPSNR方面重建的HR人脸的视觉质量比较。 |

图8.训练期间重建的HR人脸的视觉质量和识别准确性的变化。

我们方法的性能大大高于其他配置。总之，我们可以得出结论，SRnet的使用和培训可以提高我们提出的方法体系结构的识别性能，特别是对于分辨率非常低的探针图像。

## F.重建的HR面孔评估

尽管有其他耦合映射方法，我们提出的方法也可以从低分辨率的人脸重建高分辨率的人脸。在本实验中，我们旨在根据高分辨率的面部重建评估我们的方法。在这里，我们再次将我们的方法的性能与图5中介绍的三种配置进行比较，以重建人脸图像的视觉质量为依据。本节中使用的低分辨率图像的大小为24×24像素。图6示出了通过每种方法重建的面部图像的一些示例。为了比较这四种方法的视觉增强效果，使用了峰值信噪比（PSNR），结构相似性指数（SSIM）和加权峰值信噪比（WPSNR [50]）度量。如图7所示，当SRnet与FECNN网络分离时，重建的人脸图像具有最佳的视觉质量，其次是稀疏编码。我们的方法在这些结果中排在第三位，但是与前两种方法相比，我们的方法重建的人脸图像之间的差异很小。如第IV-E节所述，与其他配置相比，我们提出的方法的识别精度要好得多。这表明在我们提出的方法中，为了获得更好的识别性能，超分辨脸部图像的视觉质量受到了损害。有趣的一点是，与其他三种竞争方法相比，我们的方法的PSNR和SSIM的方差更高。这表明，在某些情况下，例如图6中的前两个示例（左侧），视觉质量得到了改善，而在另一些情况下，在其他情况下，其他两个示例的视觉质量却下降了。换句话说，SRnet的变化已朝着最终帮助识别性能的方向发展，而不一定是在视觉增强方面。图8比较了四种方法在训练过程中重建的面部图像的视觉质量和识别精度的变化。在培训阶段结束时可以看到，

V. 结束语

在本文中，我们提出了一种新颖的耦合映射方法，用于使用深度卷积神经网络识别低分辨率人脸图像。我们方法的主要思想是使用两个DCNN将低分辨率探针和高分辨率画廊人脸图像转换为一个公共空间，其中所有人脸属于同一个人的距离比人脸之间属于不同人物的距离更近。与最新的耦合映射方法（CLPM [23]，NMCF [24]，MDS [16]）和超分辨率方法（DSR [10]）相比，我们提出的方法证明了识别精度的显着提高。我们提出的方法在表情，照度和年龄的变化方面显示出显着的改进和鲁棒性。6×6 像素。我们提出的方法还提供了HR图像重建，其视觉质量可与最新的超分辨率方法相媲美。我们训练后的模型所需的空间比传统的针对人脸识别的深度卷积神经网络（如VGGnet）要少得多，因此我们提出的低分辨率人脸识别方法适用于具有常规内存的系统。