

**中国移动专利申请**

**技术交底书**

|  |  |
| --- | --- |
| **公司编号** | 由集团公司专利管理处填写 |
| **发明名称** | 基于迁移学习的人脸对齐非对齐融合的人脸伪装识别算法 |
| **申报单位** | China Mobile Technology (USA) Inc. |
| **申报类型** | 发明,实用新型 |
| **发明人** | Weisen Pan, Jian Li, Lisa Gao |
| **技术联系人** | Weisen Pan  weisenpan@chinamobile.com, +1-4082730088 |

|  |
| --- |
| **注意事项**  1．技术联系人应为深入了解本申请提案技术方案的技术人员，如交底书撰写人，负责向专利审核人员和代理人解释技术细节、修改交底书、审核申请文件等工作, 请务必填全技术联系人的姓名、E-mail、手机。  2．请按照集团公司提供的本技术交底书模板逐项填写，除交底书第八部分为可选项外，其他均为必须填写的内容。填写不全的专利申请提案，集团公司不予立案。  3．专利申请不要求已具体实现或实施，形成完整的技术方案即可提交申请，特别是需要向合作方公开、向标准提案或以其他方式公开的重要技术构思应在公开前尽早申请。  4．技术交底书文件命名要求：发明名称＋短横线（半角）＋交底书＋版本号，例：一种短消息群发方法-交底书v1.doc |

**中国移动通信集团公司**

# 摘要

基于深度学习的人脸识别目前已经取得了显著的进步，然后“伪装人脸”，不论是本人伪装后无法识别，或者他们伪装为本人，都是一项具有挑战的任务。本文基于2018年的DFW位置数据集进行伪装人脸的算法设计。在本提案中，本提案提出了一种双阶段双DCNN网络的方法。 具体来说，就是在训练阶段本提案使用通用的数据训练深度卷积网络（DCNN）用于人脸的识别。这个训练包括了人脸对齐网络DCNN-aligned和人脸非对齐网络DCNN-unaligned两个网络。 接着，本提案使用DFW数据集中的训练集，结合上面的DCNN的训练结果，使用主成分析PCA找到最近的迁移矩阵。最后实验对比了单阶段、单DCNN网络，以及当前比较前沿的其他算法，结果证明的本提案的算法是最佳的。

# 一、发明名称

基于迁移学习的人脸对齐非对齐融合的人脸伪装识别算法

# 二、技术领域

人工智能、深度学习、 人脸识别、真假脸、伪人脸

# 三、现有技术的技术方案

从2002年至今，已经开发了几种和面部伪装相关的技术。Ramanathan等人[2]使用PCA和Mahalanobis余弦距离来检测伪装的人脸。Singh等人[4]使用 2D logpolar Gabor特征来准确的验证具有伪装的人脸。其他人也提出了涉及PCA[3]或者各种纹理描述[8, 12]的其他几种方法。比起前沿的方法[9]是使用面部分类为生物特征和非生物特征。最近，深度学习已经成功应用于人脸识别。2017年Singh等人 使用深度学习的特征去对人脸的伪装进行分类，但是并没有解决人脸伪装识别的问题。相关用伪装变化检测面部和面部识别伪装的已发布的部分算法的总结见表1。

表1：用伪装变化检测面部和面部识别伪装的已发布的部分算法的总结

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 年份 | 作者 | Face Disguise Classification | Face Recognition with Disguise Variations | 使用的数据库 |
| 2002 | Martinez et al. [1] | No | Yes | AR |
| 2004 | Ramanathan et al. [2] | Yes | Yes | National Geographic, AR |
| 2005 | Kim et al. [3] | No | Yes | AR, FERET |
| 2009 | Singh et al. [4] | No | Yes | AR, Private DB |
| 2009 | Wright et al. [5] | No | Yes | AR, Yale |
| 2010 | Yang and Zhang [6] | No | Yes | AR, Yale |
| 2011 | Min et al. [7] | Yes | Yes | AR |
| 2013 | Dhamecha et al. [8] | Yes | Yes | I2BVSD |
| 2014 | Dhamecha et al. [9] | Yes | Yes | ID V1 |
| 2017 | Singh et al. [10] | Yes | No | Simple and Complex Face Disguise DBs |
| 2018 | Ours | No | Yes | DFW, VGGFace2，UMD FACES，MS1M，CelebA,CASIA-WebFace |

# 四、现有技术的缺点及本申请提案要解决的技术问题

人脸识别是计算机视觉中的经典问题。 通用的人脸识别算法旨再解决各种人脸识别的问题，尤其是基于深度卷积网络（DCNN）在这项任务中取得的巨大成功。然后，这些软件对伪装人脸的数据（比如DFW数据集）的训练不能取得很好的训练结果。

伪装人脸识别侧重于识别伪装后的人脸和模仿假冒者的身份。但是这个一共研究主题的文章都是有限的。如表1中有的方案也提出了融合卷积网络来识别面部信息，有的也提出了一个类似的数据集。但是这些方法和数据都是受到它自己的方案的限制，不适合推广。另外，DFW是提出了伪装的人脸识别数据集，但是总的数据集还是有限。本提案结合了通用数据集和伪人脸数据集DFW，进行了伪人脸识别的研究。

# 本申请提案的技术方案的详细阐述

本提案使用了一个2018年提出的一个新的伪装人脸数据集Disguised Faces in the Wild (DFW) 进行了伪装人脸识别的算法设计和识别（见图1）。DFW数据集中包括至少1000个人的至少11000张图片，这些数据集主要是从互联网手机的，因此涵盖了各种个装配件或者其他的帮助伪装的细节。

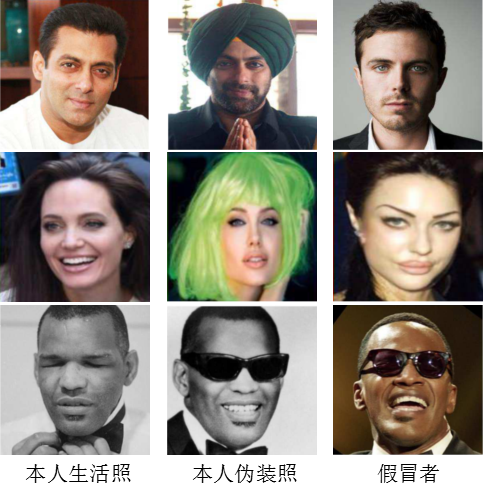


图1. DFW数据库中本人照、伪装照和假冒者的示例

在通用人脸识别训练(generic face recognition training)阶段, 本提案的数据源是几个公共数据集的合并的大数据。这个大数据包括VGGFace2，UMD FACES，MS1M，CelebA和CASIA-WebFace这5个经典的人脸数据集，一共92748个人的760万左右的图片数据（见图2）。

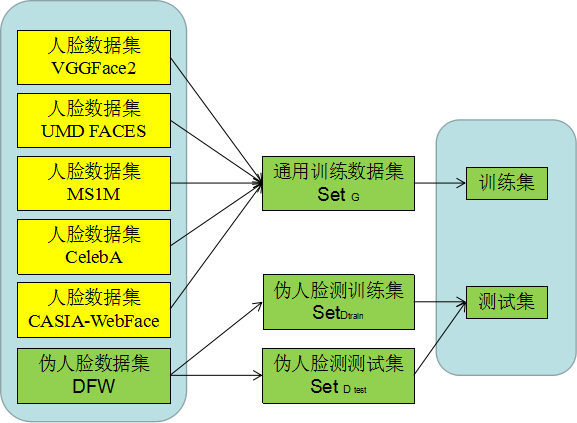


图2 本提案使用的几个数据集

## 整体框架

本提案的整体算法流程见图3。它包含了用于人脸身份特征提取的两个DCNN网络(对齐网络 DCNN-aligned和非对齐网络DCNN-unaligned), 以及一个伪装人脸适应的迁移矩阵Wselect。

对于训练过程， 首先训练两个DCNN网络用于人脸识别，然后使用主成分析(PCA)找到伪装人脸识别的适应迁移矩阵。 具体来说，在自适应的训练步骤中，根据DFW伪人脸最大样本方差的主要成分定义PCA子空间，将由DCNN提取的身份特征投影到这个PCA空间上。然后我们根据DFW训练集的经验性能选择维数。训练算法A1和测试算法A2见图3。

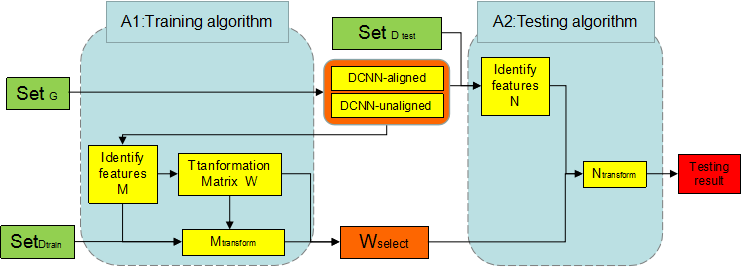


图3 整体流程示意图

## 训练阶段的方案

在人脸训练中本提案使用了64层的ResNet卷积神经网络，并使用AM-softmax做为人脸训练集的损失函数。本提案分别使用人脸对齐和未对齐的训练两个DCNN网络，这样可以融合人脸的各种姿势。对于任何一个给定的人脸，本提案均通过两个DCNNs综合提取其特征。对于DCNN-Aligned和DCNN-Unaligned的这两个深度网络模型，本提案设置的训练迭代次数是10万次。

本提案使用通用数据集SetG按照上述方法训练了两个人脸识别的DCNN深度网络DCNN-aligned和DCNN-unaligned（见图3）。然后本提案使用这两个DCNN网络提取伪装人脸数据训练集SetDtrain的身份特征(identity features) 矩阵M。接着使用主成分分析(PCA)找到适合伪装人脸适配器的迁移矩阵(transformation matrix) W。结合DFW训练集最大样本方差，使用主成分分析(PCA)得到新的迁移矩阵(transformation matrix) Mtransform。基于 Mtransform的结果在W中选择维数，最后得到Wselect。

## 测试阶段的方案

在测试阶段，对于给人的人脸图像，本提案使用两个DCNN来提取身份特征矩阵N。 然后结合Wselect，得到迁移矩阵(transformation matrix) N，最终得到表示身份的测试结果（见图3）。

## 消融实验结果

在本提案中，DCNN网络的训练方法可以是只有一个对齐网络，或者一个非对齐网络，或者是合并两个网络的这三种训练方法（见图4）。

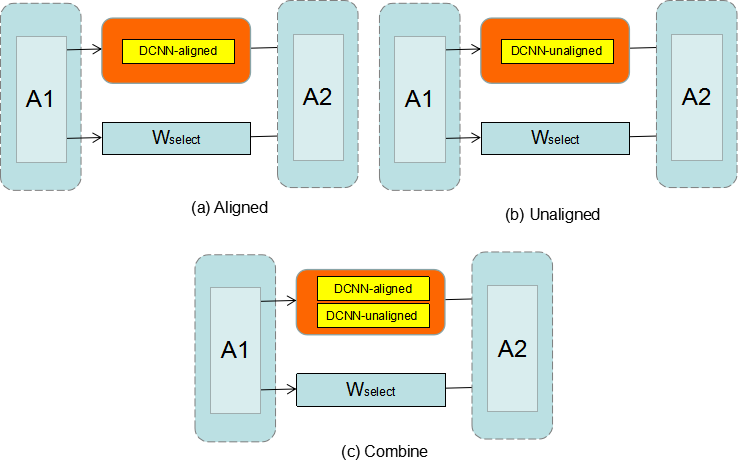
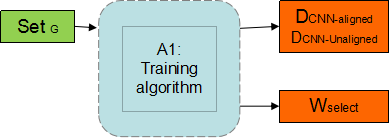
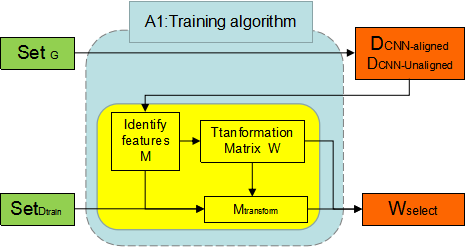


图4 DCNN的人脸对齐 、或非对齐的三种不同的训练方法

另外，DCNN网络根据是不是需要加入伪人脸数据集DFW的部分数据进行特别训练，可以划分为单阶段和双阶段发训练方法（见图5）。

1. One-stage (b)Two-stage

图5 DCNN的单阶段 、或非双阶段的两种不同的训练方法

本提案设置了记组实验测试哪种DCNN网络效果最好（见图6）。 方法M1是单阶段对齐网络，方法M2是单阶段非对齐网络，方法M3是单阶段双网络，方法M4是双阶段双网络。根据这个图6和表2中的测试结果，训练方法M1，M2,M3,M4的测试结果分别是0.842,0.847,0.849和0.889。可见是双阶段双DCNN网络的效果最优。

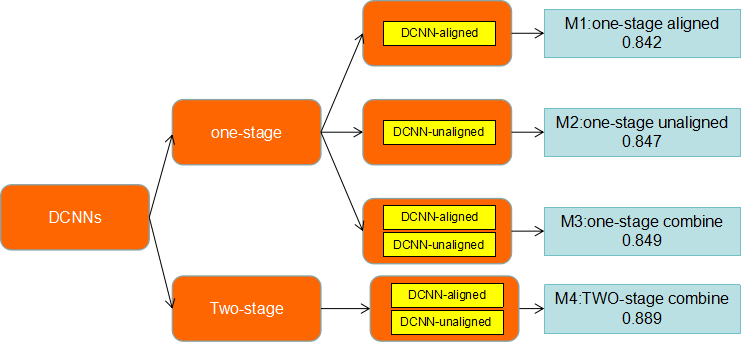


图6 几种不同的训练方法的测试示意图

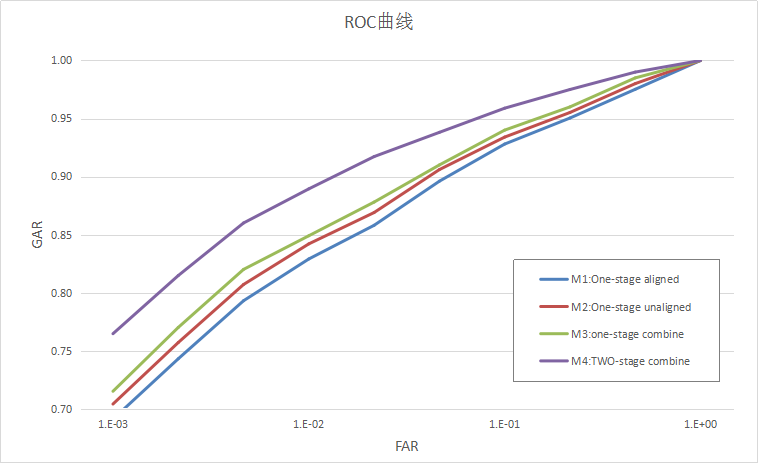


图6 评估不同训练方法的测试结果

表2： 评估不同训练方法的测试结果

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Methods** | **M1:**  **One-stage aligned** | **M2:**  **One-stage unaligned** | **M3:**  **One-stage combine** | **M4:**  **TWO-stage combine** |
| GAR@  FAR=0.1% | 0.691 | 0.704 | 0.715 | **0.765** |
| GAR@  FAR=1% | 0.842 | 0.847 | 0.849 | **0.889** |

# 本申请提案的关键点和欲保护点

本提案的主要贡献是：

1. 提出了一个双阶段训练模型。这个模型的本质就是把通用的数据训练的DCNN网络的特性转移到伪人脸数据上，把那个进行提取相关的特征地图，最后用来验证具有伪装的人脸。

1. 设置了双DCNN网络。本提案训练了两个DCNN网络，人脸对齐网络DCNN-aligned和非人脸对齐网络DCNN-unaligned，这样的训练结果就不受限于不同姿势下的人脸识别。实验表明，这种双网络方法比任何一个单网络方法更好。

# 与第三条中最接近的现有技术相比，本申请提案有何技术优点

1. 使用了2018年提出了伪人脸数据集DFW。

2. 本提案提出了一种基于DCNN的伪装人脸的双阶段训练方法， 该方法利用了有限的伪装训练数据进行伪装人脸的识别。

3. 本提案提出发方法是基于伪装人脸数据集DFW的伪人脸识别中体现了比他人更好的实验结果。

# 八、其他有助于理解本申请提案的技术资料

现有的技术方案参考文献：

[1]A. M. Mart´ınez, “Recognizing imprecisely localized, partially occluded, and expression variant faces from a single sample per class,” IEEE Transactions on Pattern analysis and Machine Intelligence, vol. 24, no. 6, pp. 748–763, 2002.

[2]N. Ramanathan, R. Chellappa, and A. R. Chowdhury, “Facial similarity across age, disguise, illumination and pose,” in IEEE International Conference on Image Processing, vol. 3,2004, pp. 1999–2002. 1, 2

[3]J. Kim, Y. Sung, S. M. Yoon, and B. G. Park, “A new video surveillance system employing occluded face detection,” in International Conference on Industrial, Engineer ing and Other Applications of Applied Intelligent Systems, 2005, pp. 65–68. 1, 2

[4]R. Singh, M. Vatsa, and A. Noore, “Face recognition with disguise and single gallery images,” Image and Vision Computing, vol. 27, no. 3, pp. 245–257, 2009. 1, 2

[5]J. Wright, A. Y. Yang, A. Ganesh, S. S. Sastry, and Y. Ma, “Robust face recognition via sparse representation,” IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 31, no. 2, pp. 210–227, 2009. 2, 6

[6]M. Yang and L. Zhang, “Gabor feature based sparse representation for face recognition with gabor occlusion dictionary,” in European conference on computer vision, 2010, pp. 448–461. 2

[7]R. Min, A. Hadid, and J.-L. Dugelay, “Improving the recognition of faces occluded by facial accessories,” in IEEE International Conference on Automatic Face & Gesture Recognition and Workshops, 2011, pp. 442–447. 1, 2

[8]T. I. Dhamecha, A. Nigam, R. Singh, and M. Vatsa, “Disguise detection and face recognition in visible and thermal spectrums,” in IEEE International Conference on Biometrics, 2013, pp. 1–8. 1, 2, 4

[9]T. I. Dhamecha, R. Singh, M. Vatsa, and A. Kumar, “Recognizing disguised faces: Human and machine evaluation,” PLOS ONE, vol. 9, no. 7, 2014. 1, 2, 4, 5, 6, 7

[10] A. Singh, D. Patil, G. M. Reddy, and S. Omkar, “Disguised face identification (DFI) with facial keypoints using spatial fusion convolutional network,” arXiv preprint arXiv:1708.09317, 2017. 2

本提案的实验结果中对比参考的文献：

[11]Parkhi, O. M., Vedaldi, A., & Zisserman, A. (2015, September). Deep face recognition. In BMVC (Vol. 1, No. 3, p. 6).

[12]Kohli, N., Yadav, D., & Noore, A. (2018). Face Verification with Disguise Variations via Deep Disguise Recognizer. In CVPR Workshop on Disguised Faces in the Wild (Vol. 4).

[13]Wen, Y., Zhang, K., Li, Z., & Qiao, Y. (2016, October). A discriminative feature learning approach for deep face recognition. In European Conference on Computer Vision (pp. 499-515). Springer, Cham.

[14]Liu, W., Wen, Y., Yu, Z., Li, M., Raj, B., & Song, L. (2017, July). Sphereface: Deep hypersphere embedding for face recognition. In The IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (Vol. 1, p. 1).

[15]Wang, F., Cheng, J., Liu, W., & Liu, H. (2018). Additive margin softmax for face verification. IEEE Signal Processing Letters, 25(7), 926-930.

[16]Smirnov, E., Melnikov, A., Oleinik, A., Ivanova, E., Kalinovskiy, I., & Lukyanets, E. (2018). Hard Example Mining with Auxiliary Embeddings. In CVPR Workshop on Disguised Faces in the Wild (Vol. 4).