

基于面部特征点运动的活体识别方法

王宇龙, 刘开元

(广州大学计算机科学与教育软件学院, 广东 广州 510006)

摘要: 提出一种应用于手机移动端基于深度学习反照片及反视频的生物识别方法, 该方法利用面部动作数据集训练一个 LSTM 网络, 通过用户配合输入一段随机动作顺序的视频, 对视频预处理后提取面目特征特点得到特征点数据, 最后将特征点数据放入循环神经网络判断视频是否发生了伪造攻击。实验结果表明, 该方法能够有效抵御照片攻击和视频回放攻击。

关键词: 活体识别; 面部特征点; 深度学习; LSTM

中图分类号: TP301.6

文献标识码: A

doi: 10.11959/j.issn.2096-109x.2018044

Living body recognition method based on facial feature point motion

WANG Yulong, LIU Kaiyuan

School of Computer Science and Education Software, Guangzhou University, Guangzhou 510006, China

Abstract: A kind of living body recognition method was proposed, which was applied in mobile terminal and based on the deep learning. A facial movements LSTM network was trained using data sets. When users input a random sequence of video, user's facial feature point data can be gained and whether the video forgery attacks happened will be determined by input user's facial feature point data into circulation neural network. The test data shows that the proposed method can be protected effectively from photograph attack and video replay attack.

Key words: living body recognition, facial feature point, deep learning, LSTM

1 引言

生物识别技术是运用声学、光学、生物传感器等, 采集人体的身体特征和行为特征, 运用计算机处理相关数据, 对个人身份进行辨别的技术^[1]。人脸识别技术作为生物识别技术的一个分支, 通过识别人脸生物特征, 从而区分本体与伪造样本, 抵御伪造攻击。现在智能穿戴设备已普及, 绝大

部分智能穿戴设备(如手机)的信息采集能力较弱, 除了图像和音频外无法获得其他信息, 因此人脸识别技术是目前最理想的手机移动端生物识别技术, 构建一个能够抵抗伪造攻击的面部识别系统对手机移动端人脸检测至关重要。

人脸识别系统面临 3 种欺骗攻击: 使用合法用户的照片、使用合法用户的视频、使用合法用户的三维模型或面具头套。其中, 照片欺骗和视

收稿日期: 2018-03-27; 修回日期: 2018-05-04

通信作者: 王宇龙, 1836925420@qq.com

基金项目: 国家自然科学基金资助项目(No.61772007); 广东省科学技术规划重点基金资助项目(No.2016B010124014)

Foundation Items: The National Natural Science Foundation of China (No. 61772007), Science and Technology Planning Key Project of Guangdong Province(No. 2016B010124014)

频欺骗是攻击人脸识别系统最频繁的方法。本文着重研究对照片欺骗和视频欺骗的防御。

近年来, 人脸识别活体检测领域迅速发展。Olga 等^[2]认为应将活体检测技术系统分为运动信息分析、纹理信息分析、活体部位分析 3 种, 并讨论了基于真伪图像存在的非刚性运动、噪声差异、人脸背景依赖等特性形成的分类器性能。在纹理信息分析方向上, Galbally^[3]提出一种基于图像纹理质量评价的方法来增强生物特征识别的安全性, 该方法只需要一张图片就可以区别真伪, 但对采样摄像头的分辨率要求很高, 因此该方法的应用场景有一定的限制; 在运动信息分析方向上, Pinto 等^[4]针对视频回放攻击提出一种基于 visual rhythm analysis 的活体检测方法, 利用对傅里叶变换后的视频计算垂直和水平的视觉节奏, 利用 3 种特征(灰度共生矩阵 GCLM、LBP、HOG)对 visual rhythm 表征和降维, 使用 SVM 分类器和 PLS(偏最小二乘)来辨别视频真伪, 但由于采样设备限制、维护成本较高等缺点, 很难大面积推广。在活体部位分析中, Kim 等^[5]提出一种基于图片失真属性分析的检测方法, 但由于需要采集 4 种属性并且建立数据库比对, 需处理的数据量相对较大, 不便于在手机端推行。此外, 还有大量学者针对活体识别技术提出了见解^[6-12], 这些方法与之前的方法相比具有更高的安全性, 可以更好地抵御伪造攻击。但这些方法对信息采集装置的要求很高, 兼容性不好, 故难以大面积在手机移动端推广。

本文提出一种环境需求低、欺骗代价高、便捷易推广、可用于手机移动端的单摄像头人脸识别方法。该方法基于面部动作特征点数据库训练的 LSTM(long short-term memory)网络模型, 通过对正负样本的学习后 LSTM 网络模型能够识别面部动作以及不同动作之间的切换是否来自照片或者视频伪造, 并且通过面部动作的随机排序使攻击者无法针对一个固定的动作序列进行视频伪造。

2 算法与模型

2.1 人脸检测与对齐及其算法分析

2.1.1 人脸检测与对齐

由于 LSTM 神经网络本身不具备空间信息感

知能力, 因此如果直接将人脸图像作为 LSTM 神经网络的输入数据, 则必须将其拉升为一维数据, 这样势必会丢失输入数据作为图像的空间信息。为了保证人脸图像的空间信息, 本文通过人脸对齐算法将人脸图像转化为带有空间信息的人脸特征点向量, 这种方法虽然丢失了许多面部细节, 但是保留了人脸的面部轮廓和表情特征, 而这正是本文所需要的数据。

本文使用 HOG+CNN 算法^[13]进行人脸识别, 以及使用 Kazemi 等^[14]发表于 CVPR 2014 上的 ERT(ensemble of regression tress)算法进行人脸特征点识别, ERT 算法是一种基于回归树的人脸特征点识别算法, 这种方法通过建立一个级联的梯度提升决策树(GBDT)使人脸形状从当前形状一步一步回归到真实形状(如图 1 所示, 其中 T 为弱回归器迭代的次数)。

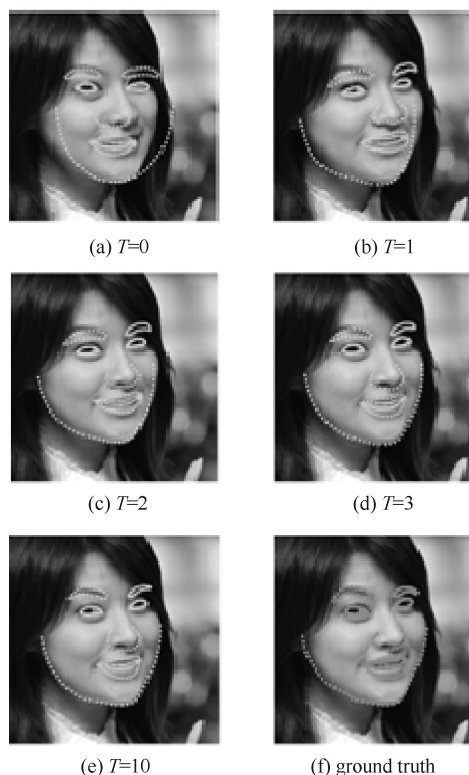


图1 人脸对齐流程

2.1.2 算法分析

1) 级联回归器

假设有训练集图像集合 $I = \{I_1, I_2, \dots, I_n\}$ 和对应的面部关键点模型集合 $S = \{S_1, S_2, \dots, S_n\}$, 将数据初始化为 $(I_{\pi_i}, \hat{S}_i^{(0)}, \Delta S_i^{(0)})$, 其中

$$\pi_i \in \{1, \dots, n\}$$

$$\hat{s}_i^{(0)} \in \{S_1, \dots, S_n\}$$

$$\Delta S_i^{(0)} = S_{\pi_i} - \hat{s}_i^{(0)}$$

π_i 为 $1 \sim n$ 中的任意值, I_{π_i} 表示图像集合 I 中下标为 π_i 的图像, $\hat{s}_i^{(0)}$ 表示除 S_{π_i} 外的任意关键点模型, $\Delta S_i^{(0)}$ 表示图像 I_{π_i} 的关键点模型与初始化模型 $\hat{s}_i^{(0)}$ 之间的残差。

级联回归器的公式如下。

$$\hat{s}_i^{(t+1)} \in \{\hat{s}_i^{(t)} + r_t(I_{\pi_i}, \hat{s}_i^{(t)})\}$$

$$\Delta S_i^{(t+1)} = S_{\pi_i} - \hat{s}_i^{(t+1)}$$

其中, $r \in R = \{r_1, r_2, \dots, r_t, \dots, r_T\}$, r 表示单个 GBDT, R 表示包含 T 个 GBDT 的回归器集合 R , t 为当前 GBDT 的编号 (下标)。级联回归的思想是通过不断迭代弱回归器使预测值与真实值的残差不断缩小, 最终得到一个较为精确的预测值。

2) 级联回归中的单 GBDT 回归树

设单 GBDT 回归树 r 中包含 K 棵弱回归树 g 和学习率 $v \in [0, 1]$, 则有公式如下。

$$r = f_K(I_{\pi_i}, \hat{s}_i^{(t)})$$

$$f_K(I_{\pi_i}, \hat{s}_i^{(t)}) =$$

$$f_k \begin{cases} \arg \min_{\gamma \in R^{2p}} \sum_{i=0}^N |\Delta s_i^{(t)} - \gamma|, k=0 \\ f_{k-1}(I_{\pi_i}, \hat{s}_i^{(t)}) + v g(I_{\pi_i}, \hat{s}_i^{(t)}), k=1, 2, \dots, K \end{cases}$$

其中, 当 $k=0$ 时, 求解一组残差向量 γ 使 $\sum_{i=0}^N |\Delta s_i^{(t)} - \gamma|$ 最小; 当 $k>0$ 时, 通过迭代弱回归树函数 g 求解残差和向量 γ 作为最终输出, 算法 1 流程如下。

算法 1 弱回归器 r_t

Training data $\{(I_{\pi_i}, \hat{s}_i^{(t)}, \Delta S_i^{(t)})\}_i^N \{I_{\pi_i}, \hat{s}_i^{(t)}\}$

Initialise, $k=0$:

$$f_k = \arg \min_{\gamma \in R^{2p}} \sum_{i=0}^N |\Delta s_i^{(t)} - \gamma|$$

For $k=1, \dots, K$

Set for $i=1, \dots, N$

$$r_{ik} = \Delta S_i^{(t)} - f_{k-1}(I_{\pi_i}, \hat{s}_i^{(t)})$$

Fit a regression tree $g_k(I, \hat{s}_i^{(t)})$

Update

$$f_k = f_{k-1}(I, \hat{s}^{(t)}) + v g(I, \hat{s}^{(t)})$$

$$\text{Output } r_t(I, \hat{s}^{(t)}) = f_k$$

2.2 基于面部动作序列的活体检测方法

基于面部动作序列的活体检测方法概述如下。

- 1) 系统初始化, 载入训练好的 LSTM 网络模型。
- 2) 对输入的视频流进行人脸检测。
- 3) 检测到人脸后, 产生一个随机的面部动作序列并且指导用户按顺序执行面部动作。
- 4) 计算得到用户的人脸特征点, 并输入 LSTM 网络模型进行面部动作判断。
- 5) 检测过程中如果某一帧人脸消失无法检测, 则回到步骤 2) 重新开始检测。

6) 得到 LSTM 网络模型的面部动作判别结果序列, 比对结果序列是否与步骤 3) 中产生的随机面部动作序列相同。

2.3 基于特征点数据的动作识别及其算法分析

2.3.1 基于特征点数据的动作识别

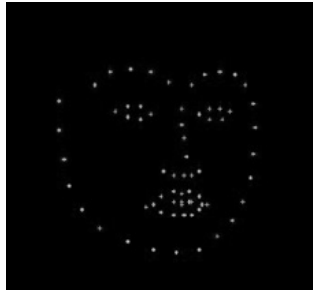
在通常情况下, 人脸识别系统面临着 3 种欺骗手段: 1) 使用合法用户的照片; 2) 使用合法用户的视频; 3) 使用合法用户的三维模型或面具头套。

其中, 三维建模或面具头套是最具威胁的伪造方式, 一个足够精细的三维建模能够在程序的控制下做出任何伪造者想要的动作或者表情, 这对识别系统是致命的, 但这也要求伪造者获得合法用户几乎所有的面部细节, 除非合法用户主动暴露否则几乎无法获得, 因此本文主要针对前 2 种欺骗手段, 通过运用 LSTM 神经网络对人脸特征点的分布以及运动轨迹进行分类, 从而区分用户的合法性。

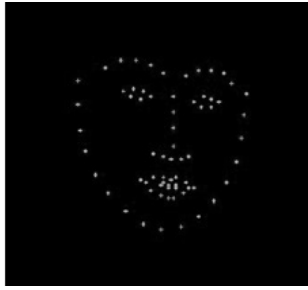
通过实验对比, 伪造者尝试通过使用合法用户的照片进行欺骗时的面部特征点位置和运动轨迹与正常的合法用户的面部特征点位置是不同的。例如, 图 2 是摇头动作时的侧脸图, 其中, 图 2(a) 为合法用户的面部特征点分布, 图 2(b) 为照片欺骗时的面部特征点分布。

当伪造者尝试通过使用合法用户的视频进行欺骗时, 由于系统生成的动作序列是随机的, 因此伪造者无法在动作序列未知的情况下获得合法用户的完整合法视频, 只能通过将包含合法用户面部动作的视频进行合成来获得正确序列的伪造

视频。由于伪造视频是合成的，所以不同面部动作时间的衔接动作会十分生硬甚至没有衔接动作，而合法的用户视频不同面部动作之间的衔接会更为自然、平滑，这些区别同样也会在特征点序列中展现出来。



(a) 合法特征点图



(b) 伪造特征点图

图2 合法/伪造特征点图

本文通过 LSTM 神经网络学习点头、摇头、微笑 3 个面部动作的运动轨迹特征和不同面部动作之间衔接时的运动轨迹特征，从而正确区分合法的用户视频和照片欺骗或者合成视频重放欺骗。

2.3.2 算法分析

RNN (recurrent neural network)^[15]是一种专门设计用于处理序列问题的网络结构，能让输入的信息持久化，其网络结构如图 3 所示，RNN 通过循环结构将信息沿序列传递下去，使任意序列节点 RNN 的输出都是与以往节点的信息紧密相关的。

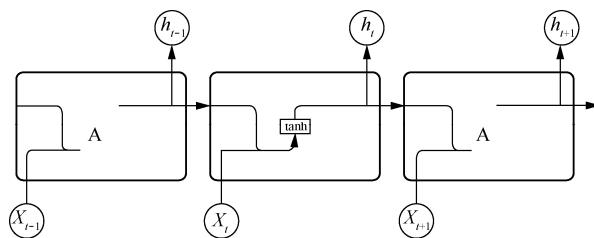


图3 RNN 结构

随着研究的不断深入，研究人员发现随着输入序列长度的变化，RNN 的效果会越来越差，出现这个问题的关键是 RNN 在处理输入间隔不同的数据信息时并不是“一视同仁”的，随着这个间隔的不断增大，RNN 的学习能力会不断递减，最终导致与 RNN 间隔最短的信息权重最大，间隔越大的信息权重越小甚至没有，从而出现长期依赖 (long-term dependencies) 问题。

为了解决长期依赖问题，Hochreiter 与 Schmidhuber^[16]于 1997 年提出了一种经过改良的 RNN 模型 LSTM 网络，这种网络结构通过增加门 (gates) 结构和状态信息来避免长期依赖问题，其网络结构如图 4 所示。

每个节点中包含 3 个门结构来控制节点对状态信息的影响，门是一种让信息选择式通过的方法。它包含一个 sigmoid 神经网络层和一个 pointwise 乘法操作，sigmoid 层输出 0~1 的数值，描述当前信息的通过百分比，0 代表“不许任何信息通过”，1 就指“允许所有信息通过”。状态信息 C 则会穿过循环神经网络上的所有节点，每个输入序列在经过处理后得到的信息不再只传给下一个输入序列，而是经过加权后传入状态信息中。状态信息就像一条光纤，保证了不同节点的信息传递不受间隔距离的影响。LSTM 增加的门包含遗忘门 (forget gates)、传入门 (input gate) 以及输出门 (output gate)，分别控制当前时间点 t 下数据的输入、修改和输出。

遗忘门的算法公式为

$$f_t = \sigma(W_f[h_{t-1}, x_t] + b_f)$$

其中， σ 表示 sigmoid 激活函数， W_f 表示待学习系数， h_t 表示 t 时间时，LSTM 节点的输出， x_t 表示 t 时间时的输入数据， b_f 表示偏转因子。遗忘门的主要作用是决定上一节点传来信息的通过率，通过系数 W_f ，遗忘门能够去除不需要传递的冗余信息。

传入门的算法公式为

$$I_t = \sigma(W_i[h_{t-1}, x_t] + b_i)$$

$$\tilde{C}_t = \tanh(W_c[h_{t-1}, x_t] + b_c)$$

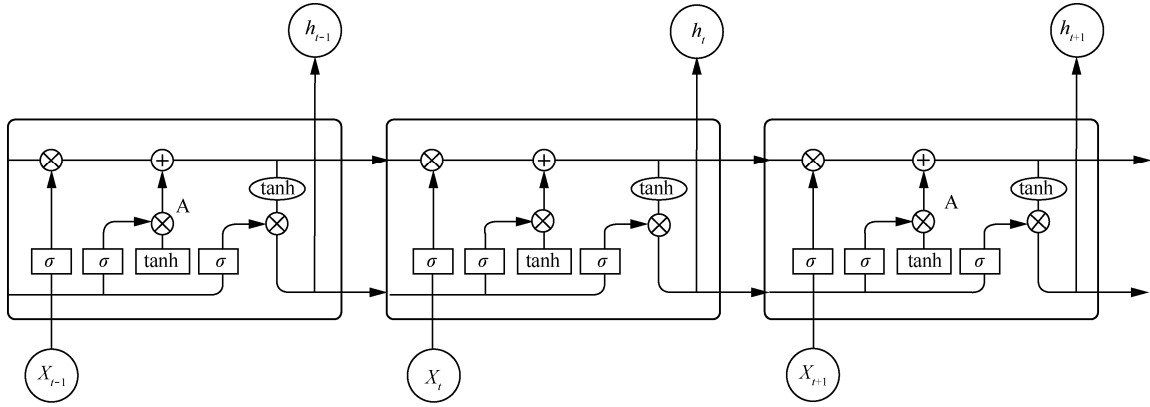


图4 LSTM网络结构

$$C_t = f_t C_{t-1} + I_t \tilde{C}_t$$

不难发现,输入门包含3部分,其中,第一部分 I_t 用于判断上一节点的数据中哪部分需要更新;第二部分 $\tilde{C}_t$ 用于提取当前节点的输入 x_t 中用于更新的内容;第三部分 C_t 表示经过计算后需要汇入状态信息 C 的数据。也就是说, I_t 用于标注句子中哪部分需要更新, $\tilde{C}_t$ 用提取更新的内容, C_t 则联合 $\tilde{C}_t$ 、 I_t 和 f_t 对来自上一节点的状态信息 C_{t-1} 进行更新。

输出门的算法公式为

$$O_t = \sigma(W_o[h_{t-1}, x_t] + b_o)$$

$$h_t = O_t \tanh(C_t)$$

由于输入门仅对上一节点的状态信息中需要更新的部分进行了更换,但是当前时间节点的信息中仍包含之前状态信息中没有的信息。例如,上一节点传入了“你好”,输入门仅对已存在的部分进行更新,如将主语“你”更新为“你们”,但是并没有添加原句子中不存在的部分。因此,输出门的作用在于从当前时间点的输入 x_t 中抽取需要传递的新信息并与当前的状态信息 C_t 进行整合,并与样本标注 l 进行比对,通过最小化交叉熵对模型进行训练。

$$y = h_t$$

$$y' = e^y$$

$$y'' = \frac{y'_j}{\sum_{j=1}^n y'_j}$$

$$\text{loss} = \arg \min \left(-\sum_{i=1}^N l_i \log(y'') \right)$$

其中, y 表示 t 时间点时 LSTM 的输出, y'' 为模型输出的交叉熵函数, n 表示样本分类总数, l_i 表示当前样本的真实标注, N 表示总样本集合, $i \in [1, N]$ 。

本文通过将 ERT 算法的输出作为训练数据 $X = \{x_i, l_i\}_i^N$, 其中, l_i 表示对应数据的分类标签, 设 LSTM 模型的时间序列长度为 T' , 每个训练数据 x_i 的时间长度为 T , 则 LSTM 算法流程如下。

算法2 LSTM 回归器 $model_{LSTM}$

Training data $X = \{x_i, l_i\}_i^N$

Initialise, $t=0$:

$$f_t = \sigma(W_f[0, x_t] + b_f)$$

For $t=1, \dots, T$

Set for $i=1, \dots, N$

$$f_t^i = \sigma(W_f[h_{t-1}^i, x_t^i] + b_f)$$

$$I_t^i = \sigma(W_f[h_{t-1}^i, x_t^i] + b_i)$$

$$\tilde{C}_t^i = \tanh(W_c[h_{t-1}^i, x_t^i] + b_c)$$

$$C_t^i = f_t^i C_{t-1}^i + I_t^i \tilde{C}_t^i$$

$$O_t^i = \sigma(W_o[h_{t-1}^i, x_t^i] + b_o)$$

$$h_t^i = O_t^i \tanh(C_t^i)$$

If $t > T'$

Fit a LSTM network $model_{LSTM}(x, l)$

Update

$$y = h_t$$

$$y' = e^y$$

$$y'' = \frac{y'_j}{\sum_{j=1}^n y'_j}$$

$$\min \left(-\sum_{i=1}^N l_i \log(y'') \right)$$

Output $model_{LSTM}(x, l) = y$

3 实验

3.1 实验数据及参数

为了测试本文提出的算法对活体人脸和欺骗攻击的辨别能力, 本文采集构建了一个人脸动作视频的数据库。该数据库基于 logitech C270 采集, 每个测试样本包含 12 个采集视频和 6 个合成视频, 其中, 采集视频包含活体的点头、摇头、微笑 3 个面部动作之间相互切换的视频片段以及有照片遮挡的 3 个面部动作之间相互切换的视频片段, 而合成视频则是通过视频拼接合成的面部动作切换视频, 每个视频分辨率为 640×480 , 采样频率为 25 帧/秒, 每个视频片段时长大约 20 s, 一共采集了 8 个人, 共 144 个视频片段, 其中, 挑选了 12 个合法用户视频、12 个照片欺骗视频和 12 个合成伪造视频, 共 36 个视频片段作为测试集, 其余作为训练集。

本实验所使用的平台配置为 I7-7700K 处理器, 16 GB 内存, GTX1080 显卡, 人脸特征点检测采用 dlib 库, 该库实现了基于 ERT 算法的人脸特征检测模型; 面部动作识别以及活体检测使用的是基于 Tensorflow 平台的 LSTM 神经网络模型。在使用 GPU 加速的情况下系统的处理速度能达到 30 帧/秒, 表明本文所提出的方法可以在有实时要求的项目背景下使用。

实验模型参数设置如下。

本文的 ERT 算法模型采用的为 68 点人脸特征预训练模型, LSTM 网络为 3 层隐藏层, 每层节点数设置为 256 个, 为了避免过拟合, 所有 LSTM 网络共享一套权值; 训练过程的误差传递使用随机梯度下降 (SGD) 的方法, 学习率设置为 0.1, 对于误差的更新则采用批处理的形式, 每次由 16 个样本一起更新, 即 batch 设置

为 16 个, dropout 设置为 0.2, 状态机序列长度设置为 50, 轮回次数设置为 2 000, 选择在评估数据集上性能最好的一组参数作为训练的模型参数输出。

3.2 验证与分析

本文通过测试集中的 36 个视频片段测试了本文方法的活体检测性能, 其结果如表 1 所示。

表 1 本文方法的活体检测性能结果

视频类型	召回率	精确率	F1
照片欺骗	1	1	1
视频合成	1	0.923 0	0.959 9

实验结果显示, 在 36 个样本的测试集中只在分辨合成视频中发生了误识别, 误识别样本数为 1; 照片欺骗则全部样本都成功识别。需要注意的是, 实验的前提假设为攻击者已经获得了合法用户的照片以及合法用户部分面部动作片段, 能够在通过视频合成的方式得到正确动作序列的面部视频的情况下, 实验所用的测试样本较少, 因此在实际环境中的各项数值会更高。

众所周知, 由于神经网络模型属于黑盒算法, 因此对于实验的结果以及算法参数是难以解释的, 为了能让本文所使用的 LSTM 网络模型能够有个直观的解释, 本文计算得到了摇头动作下合法用户 (valid users) 和照片欺骗 (forge) 下不同面部区域的特征点近距离均值变化图作为参考。

图 5 显示了面部中左眼 (left eye)、左眉 (left eyebrow)、右眼 (right eye)、右眉 (right eyebrow)、鼻尖 (nose top)、鼻梁 (nose bridge)、上唇 (top lip)、下唇 (bottom lip)、下巴 (chin) 共 9 个区域的特征点距离均值变化, 其中, 横坐标为帧数, 纵坐标为当前帧特征点与上一帧对应特征点的平均距离。不难发现, 除了眼部和眉毛区域以外, 照片欺骗视频和合法用户视频的特征点距离均值的分布和振幅都有较大的差异。

同样, 本文计算得到合法用户不同面部动作之间变换时面部特征点的位移变化量如图 5 所示。

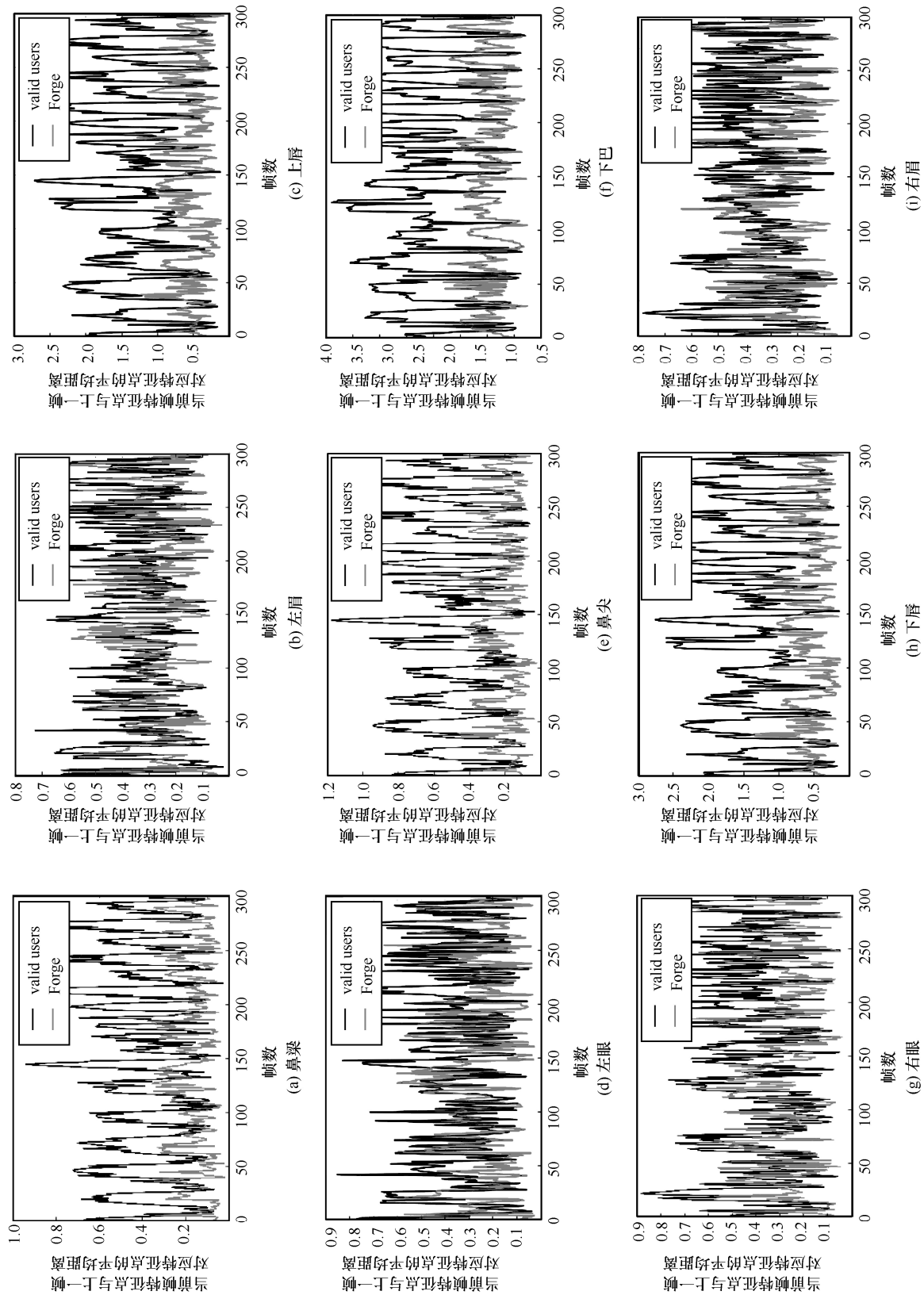


图 5 特征点变化

图6是合法用户视频以及合成视频摇头切换到点头的100帧影像片段,从图中可以看出,合成的视频在动作切换过程中特征点位移变化量会有一个非常大的峰值,然后恢复到正常状态,这是由于合成视频在2个面部动作之间过度不自然导致的。

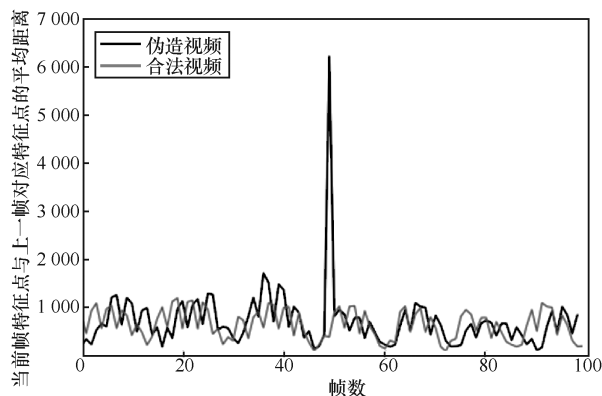


图6 合法/伪造视频用户影像片段

由图6和实验结论可知,不论是照片欺骗还是合成视频回放攻击,在面部特征点的分布和变化量等特征上都与合法用户视频有较大区别,通过LSTM神经网络进行特征学习是行之有效的。

4 结束语

本文提出了利用面部特征点和LSTM神经网络的方法进行人脸认证中的活体检测。该方法主要通过获得合法用户与伪造攻击(照片欺骗,合成视频)中人脸面部特征点的分布和运动等特征数据,利用LSTM神经网络对这些特征数据进行学习和分析判别从而区分出两者的不同以达到活体检测的效果。

活体检测作为一种反攻击措施,其性能评价通常着重为攻击行为的实时成本和破解的难易程度,因为要求一种反攻击措施做到理论上的不可破解是极难的,和密码学领域一样,所有的加密算法理论上都可以通过暴力方法破解,而这些算法能够实用化的原因在于破解成本过高。对于本文所提出的活体检测方法,有效的攻击手段为:1)合法用户高精度的三维模型或者头套;2)合法用户完整的正确面部动作序列视频。

对于第一种攻击手段,伪造者需要借助专业设备来获得合法用户完整而详尽的面部信息,这在用户不主动配合的情况下几乎是不可能的;对于第二种攻击手段,伪造者需要在系统给出随机序列后的极短时间内获得合法用户的正确面部动作序列视频,或者合成出一段难辨真假的伪造视频,前者在合法用户不主动配合的情况下同样是难以完成的,后者在短时间内也是无法完成的。

综上所述,以上2种攻击手段都需要很高的技术要求或时间成本,因此本文所提出的活体检测方法在真实的项目环境中是能够有效抵挡伪造攻击的。

参考文献:

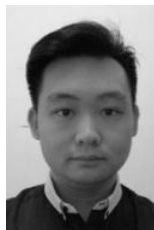
- [1] 徐子明.生物识别技术概况及其应用[J].科学技术创新, 2017, (35): 191.
- [2] XU Z M. Biometrics technology and its application[J]. Science and Technology Innovation, 2017, (35): 191.
- [3] OLGA K, NASER D. 2DFace liveness detection: an overview[C]//International Conference of the Biometrics Special Interest Group (BIOSIG). 2012: 1-12.
- [4] GALBALLY J, MARCEL S, FIERREZ J. Image quality assessment for fake biometric detection: application to iris, fingerprint, and face recognition[J]. IEEE Transactions on Image Processing, 2014, 23(2): 710-724.
- [5] PINTOA, SCHWARTZ W R, PEDRINI H, et al. Using visual rhythms for detecting video-based facial spoof attacks[J]. IEEE Transactions on Information Forensics and Security, 2015, 10(5): 1025-1038.
- [6] KIM W, SUH S, HAN J J. Face liveness detection from a single image via diffusion speed model[J]. IEEE Transactions on Image Processing, 2015, 24(8): 2456-2465.
- [7] CHINGOVSKA I, ANJOS A, MARCEL S, et al. On the effectiveness of local binary patterns in face anti-spoofing[C]//Biometrics Special Interest Group. 2012:1-7.
- [8] ANJOS A, MARCEL S. Counter-measures to photo attacks in face recognition: a public database and a baseline[C]//International Joint Conference on Biometrics. 2011:1-7.
- [9] TAN X, LI Y, LIU J, et al. Face liveness detection from a single image with sparse low rank bilinear discriminative model[C]//European Conference on Computer Vision. 2010:504-517.
- [9] ZHANG Z, YAN J, LIU S, et al. A face antispoofing database with

- diverse attacks[C]//Iap International Conference on Biometrics. 2012:26-31.
- [10] SUN L, PAN G, WU Z, et al. Blinking-based live face detection using conditional random fields[C]//International Conference on Advances in Biometrics. 2007:252-260.
- [11] BAO W, LI H, LI N, et al. A liveness detection method for face recognition based on optical flow field[C]//International Conference on Image Analysis and Signal Processing. 2009:233-236.
- [12] BHARADWAJ S, DHAMECHA T I, VATSA M, et al. Computationally efficient face spoofing detection with motion magnification[C]//IEEE Conf Comput Vis Pattern Recognit Workshops (CVPRW). 2013:105-110.
- [13] ZHANG K, ZHANG Z, LI Z et al. Joint face detection and alignment using multitask cascaded convolutional networks[J]. IEEE Signal Processing Letters, 2016, 23(10):1499-1503.
- [14] KAZEMI V, SULLIVAN J. One millisecond face alignment with an ensemble of regression trees[C]//2014 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2014: 1867-1874.
- [15] CHEN S H, HWANG S H, WANG Y R. An RNN-based prosodic information synthesizer for mandarin text-to-speech[J]. IEEE Transactions on Speech & Audio Processing, 1998, 6(3):226-239.
- [16] HOCHREITER S, SCHMIDHUBER J. Long short-term memory[J]. Neural Computation, 1997, 9(8):1735.

[作者简介]



王宇龙 (1991-), 男, 江苏徐州人, 主要研究方向为深度学习与信息安全。



刘开元 (1991-), 男, 安徽宣城人, 主要研究方向为算法优化、深度学习与图像处理。