

基于优化 LeNet-5 的近红外图像中的静默活体人脸检测

黄俊¹, 张娜娜², 章惠¹

(1. 上海海洋大学信息学院, 上海 201306; 2. 上海建桥学院信息技术学院, 上海 201306)

摘要: 针对当前交互式活体检测过程繁琐、用户体验性差的问题, 提出了一种优化 LeNet-5 和近红外图像的静默活体检测方法。首先, 采用近红外光摄像头构建了一个非活体数据集; 其次, 通过增大卷积核、增加卷积核数目、引入全局平均池化等方法对 LeNet-5 进行了优化, 构建了一个深层卷积神经网络; 最后, 将近红外人脸图片输入到模型中实现活体静默活体检测。实验结果表明, 所设计的模型在活体检测数据集上有较高的识别率, 为 99.95%, 整个静默活体检测系统的运行速度约为 18~22 帧/s, 在实际应用中鲁棒性较高。

关键词: LeNet-5; 卷积神经网络; 全局平均池化; 近红外图像; 静默活体检测

中图分类号: TP399 **文献标识码:** A **文章编号:** 1001-8891(2021)09-0845-07

Silent Live Face Detection in Near-Infrared Images Based on Optimized LeNet-5

HUANG Jun¹, ZHANG Nana², ZHANG Hui¹

(1. College of Information Technology, Shanghai Ocean University, Shanghai 201306, China;

2. College of Information Technology, Shanghai Jian Qiao University, Shanghai 201306, China)

Abstract: An improved method of silent liveness detection for LeNet-5 and near-infrared images is proposed to overcome the problem of the interactive liveness detection process and poor user experience. First, a face attack dataset was constructed using a near-infrared camera. Second, the LeNet-5 was optimized by increasing the number of convolution kernels and introducing global average pooling to construct a deep convolutional neural network. Finally, the near-infrared face image is input to the model to realize silent liveness detection. The experimental results show that the proposed model has a higher recognition rate for the liveness detection dataset, reaching 99.95%. The running speed of the silent liveness detection system is approximately 18-22 frames per second, which shows high robustness in practical applications.

Key words: LeNet-5, convolutional neural network, global average pooling, near-infrared image, silent liveness detection

0 引言

随着人脸识别技术的普及和计算机视觉技术的蓬勃发展, 人脸支付、人脸登录、人脸取件等系统已被实施应用在人们的生活中。但是, 在一些特殊场景下, 人脸识别系统极易被外来手段攻击, 如: 电子屏幕翻拍、照片打(冲)印、三维人脸模型等。因此, 为了维护人脸识别系统的安全, 保障用户的利益, 人脸活体检测技术显得相当重要。目前, 市面上及研究领域中被用的最广泛的是交互式的活体检测方法。

文献[1]提出了一个关于眼睛和嘴部动作的交互

式活体检测方法, 主要思想是计算用户眼睛区域和牙齿的 HSV (Hue, Saturation, Value) 色彩空间, 以此来判断它们的开合状态。文献[2]设计了一个关于嘴部状态和头部姿态的交互式活体检测系统, 作者利用支持向量机 (support vector machine, SVM)^[3]和面部特征点算法, 来预测头部姿态方向和嘴部开合状态, 通过随机指令要求用户做出对应动作, 以实现活体检测。文献[4]要求用户完成一些随机表情动作, 并计算连续视频帧的 SIFT (scale-invariant feature transform) 流能量值来判断表情变化。文献[5]则在虚拟柜员机 (virtual teller machine, VTM) 的相机条件下, 利用眼球色素

收稿日期: 2020-12-01; 修订日期: 2021-01-20。

作者简介: 黄俊 (1996-), 男, 浙江温州人, 硕士研究生, 主要研究方向: 图像处理、计算机视觉。E-mail: huangj_sg@163.com。

通信作者: 张娜娜 (1979-), 女, 山东莱阳人, 副教授, 硕士, 主要研究方向: 图像处理。E-mail: nanazhang2004@163.com。

基金项目: 上海市教育委员会“晨光计划”基金项目 (AASH1702)。

变化进行眨眼检测,加以背景检测、张嘴检测和微笑检测等组合命令得以实现交互式活体检测。

交互式活体检测存在着人机交互不友好、认证过程繁琐等缺点,针对其弊端,本文提出了一种基于卷积神经网络(convolutional neural network, CNN)和近红外光的静默活体检测方法,利用近红外光的成像特点及光在材质上的反射特性,直接去除了电子屏幕、相片冲印的人脸攻击;利用 CNN 自动提取图像特征的优势,对近红外光下的真人人脸图像和照片打印人脸图像进行分类,相对于可见光下的人脸图像,近红外人脸图像特征明显,更易于区分活体人脸与非活体人脸。针对上述内容,主要做了如下两个重点工作:

1) 自建非活体人脸数据集(照片打印人脸),利用近红外摄像头采集包含人脸的照片打印图像;

2) 对 LeNet-5^[6]的结构、卷积核大小、特征图数目、全连接层等部分进行研究讨论,修改并设计了一个活体分类模型,通过实验验证了该模型在活体检测方面有着较高的识别率。

1 卷积神经网络和 LeNet-5

1.1 卷积神经网络

1.1.1 卷积层

在 CNN 中,卷积层是核心,它的主要作用是提取图像特征,提升图像分类准确率。该层由若干个卷积核和激活函数组成,每个卷积核参数均通过反向传播进行不断地更新,最终提取出多个特征图。通常第一个卷积层提取的图像特征较为模糊,但随着卷积层数的增加,网络提取能力越来越强,图像特征也越来越清晰。卷积层的数学表达式^[7]如下所示:

$$x_j^l = f(\sum_{i \in m_j} x_i^{(l-1)} * k_j^l + b_j) \quad (1)$$

式中: k 为卷积核矩阵; $x_i^{(l-1)}$ 为上层输出特征图; b 为加性偏置项; $f(\cdot)$ 为激活函数,一般采用 ReLU (Rectified Linear Unit) 激活函数。

1.1.2 池化层

卷积操作虽然可以提取图像特征,但输出的特征图参数量大、信息杂,在网络训练时会严重占用计算资源,加重过拟合现象,所以,为了提高计算和运行速度,缩减模型参数,提高模型泛化能力,池化层经常作为卷积层的下一层输入单元,其数学表达式^[8]为:

$$x_j^l = f(\beta_j^l \text{down}(x_j^{l-1}) + b_j^l) \quad (2)$$

式中: $\text{down}(\cdot)$ 为下采样函数; β_j^l 为权重项; b_j^l 为偏置项。

1.1.3 全局平均池化

CNN 中的参数主要集中在全连接层中,是导致模型发生过拟合的现象的源头,全局平均池化(global average pooling, GAP)^[9]是一种能有效解决过拟合的主要方法之一。它的主要原理是计算上层特征图的像素均值,从而获得相关低维特征。相比于全连接层,它更好地保留了空间信息,降低了参数量,具有更强的鲁棒性。相关数学表达式为:

$$f_k = \text{average}(f_{i,j,k}) \quad (3)$$

式中: (i, j) 表示像素位置; k 表示通道索引。

1.1.4 输出层

输出层也叫做 SoftMax 层,该层含有一个 SoftMax 分类器,作用是归一化全连接层的输出结果,经过 SoftMax 函数的计算,会返回一组概率值,在模型推导时,概率值最高的一类可作为模型的分类结果。有关该层的数学表达式为:

$$y_i = e^{a_i} / \sum_{j=1}^n e^{a_j} \quad (4)$$

式中: a_i 为上层结构的输出结果; n 为类别数。

1.1.5 交叉熵损失

交叉熵损失反映了真实值和预测值的概率分布差异,训练中,交叉熵主要用于计算 SoftMax 输出的概率损失,值越低,模型的预测效果就越好。有关交叉熵损失的数学表达式为:

$$H(p, q) = -\sum_{i=1}^n p(x_i) \log(q(x_i)) \quad (5)$$

式中: n 为类别数, $p(x)$ 为真实分布值, $q(x)$ 为预测分布值。在分类模型的训练中,由于独热编码(One-Hot)的特殊性,损失函数可以写成如下表达式:

$$\text{Loss} = -\log(q(x_i)) \quad (6)$$

1.2 LeNet-5 模型

LeNet-5 是深度学习中较经典的 CNN 模型之一,是一种结构较为简单的 CNN,最早被用来进行手写数据集的识别。LeNet-5 主要包含 7 个结构层,其中有 3 个卷积层(C1、C3、C5)、2 个池化层(S2、S4)、1 个全连接层(F6)和 1 个输出层(Output),此外,LeNet-5 还包含一个输入层,该层的图片输入大小为 32×32 。

C1 的核大小为 5,步长为 1,输入层的图片经过该层操作之后会输出 6 个 28×28 的特征图;

S2 的核大小为 2,步长为 2, C1 输出的特征图经过该层操作之后会输出 6 个 14×14 的特征图;

C3 的核大小为 5,步长为 1, S2 输出的特征图经过该层操作之后会输出 16 个 10×10 的特征图;

S4 的核大小为 2,步长为 2, C3 输出的特征图经

过该层操作之后会输出 16 个 5×5 的特征图；

C5 的核大小为 5，步长为 1，S4 输出的特征图经过该层操作之后会输出 120 个 1×1 的特征图；

F6 是全连接层，共有 84 个神经元结点，层中每一个神经元均与 C5 中的神经元相连接；

Output 是 LeNet-5 的最后一层，因为 LeNet-5 是一个 10 分类的模型，所以该层共有 10 个神经元节点。

由于本文实现的是一个 2 分类（活体、非活体）的任务，且样本数据比手写数据集复杂，所以使用 LeNet-5 进行活体图片的分类显然是不合理的。因此，本文参考 LeNet-5 的结构与思想，在该结构上做进一步研究，以达到任务要求，以下是本文对 LeNet-5 的改进方案。

2 优化 LeNet-5 的人脸活体检测模型

在真实场景中，模型的识别准确率和实时性是评价系统是否合格的重要指标，因此，本文经过多次实验验证，在保证模型实时性较高的前提下提升了准确率。最终确定了活体检测模型的结构并将其命名为 LeNet_Liveness，其中，主要进行了以下优化：

1) 改变部分卷积核大小。常见的卷积核大小有 7×7 、 5×5 和 3×3 等，从数学角度出发， 7×7 的卷积核能提取到更多的图像细节，即卷积核越大，感受野越大，提取到的特征越多，有利于提升图像分类的精度。由于图像中的“人脸”目标远大于“数字”目标，特征复杂，因此，为了在前期不丢失太多的图像信息，本文将第一和第二卷积层的核大小设为 7，以增大感受野，扩大提取细节，后面几个卷积层的核大小保持不变，依旧为 5。

2) 增加卷积核个数。与手写数据集相比，活体检测数据集稍复杂，使用较少的卷积核无法表达图像信息，而使用较多的卷积核又会带来计算量的增加，但有助于提升图像分类的准确率，基于此考虑，本文最初在模型第一层选择 32 个卷积核，一来，模型参数不会增加太多；二来，能提取出更多的低层次信息，保留更多细节特征。

3) 加深模型结构。浅层神经网络由于复杂度较低，无法更好地提取到图像中的主要特征，使得模型分类效果欠佳，最终影响模型的识别率，本文在 LeNet-5 的结构层数上做了修改，将原来的 3 层卷积增加至 5 层，以提高网络的拟合能力。

4) GAP 替代全连接层。CNN 大部分参数集中在全连接层中，如果采用传统的全连接层，势必会造成模型参数的增大，而 GAP 属于无参结构，那么整个模型的参数就能有所降低，对于模型后期的预测速度能有所提高，除此之外，还能有效防止过拟合。

图 1 为 LeNet_Liveness 的结构图，表 1 为模型的相关结构参数。输入层为 128×128 的三通道人脸图片。C1、C2、C3、C4 和 C5 是卷积层，移动步长均为 1，其中，C1、C2 层的卷积核大小为 7×7 ，C3、C4 和 C5 层的卷积核大小为 5×5 ，除此之外，C1 输出 32 个特征图，C2 输出 64 个特征图，C3 输出 128 个特征图，C4 输出 256 个特征图，C5 输出 512 个特征图。P1、P2、P3、P4 和 P5 均为最大池化层，即对每个输入的特征图进行 2×2 的最大值运算，步长为 2，以缩减图像长宽，减小计算量。GAP 为全局平均池化层，Output 为输出层。

表 1 模型结构参数

Layer Name	Layer Type	Output Size/Strides	Kernel Size
Input	Input layer	$128 \times 128 \times 3$ / -	-
C1	Convolution	$128 \times 128 \times 32$ / 1	7
P1	Max Pooling	$64 \times 64 \times 32$ / 2	2
C2	Convolution	$64 \times 64 \times 64$ / 1	7
P2	Max Pooling	$32 \times 32 \times 64$ / 2	2
C3	Convolution	$32 \times 32 \times 128$ / 1	5
P3	Max Pooling	$16 \times 16 \times 128$ / 2	2
C4	Convolution	$16 \times 16 \times 256$ / 1	5
P4	Max Pooling	$8 \times 8 \times 256$ / 2	2
C5	Convolution	$8 \times 8 \times 512$ / 1	5
P5	Max Pooling	$4 \times 4 \times 512$ / 2	2
GAP	GAP	$1 \times 1 \times 512$ / 1	4
Softmax	Softmax	2 / -	-

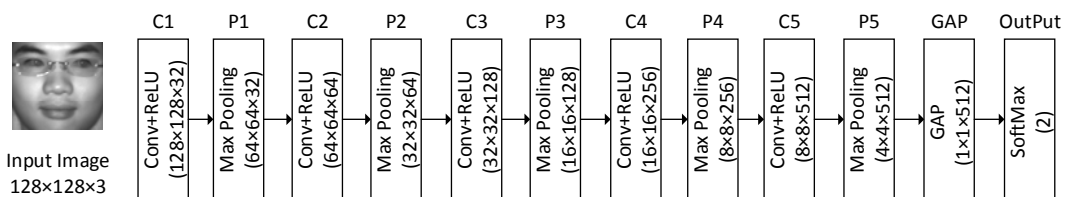


图 1 LeNet_Liveness 结构图

Fig.1 LeNet_Liveness structure diagram

3 实验与结果分析

3.1 实验环境与评估指标

实验的硬件环境为 Intel Core i5-8300H 处理器, NVIDIA GeForce GTX1050TI (4GB 显存), 内存为 8GB 大小, 深度学习框架为 TensorFlow-1.13.1, 图像处理库为 OpenCV-4.1.0, 编程语言为 Python-3.6.7, 使用 1920×1080 的近红外摄像头(带补灯光, 850 nm)采集数据。

分类模型通常使用准确率 (Accuracy) 作为其性能的评估标准, 相关计算公式如下:

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (7)$$

式中: “TP”表示正样本中预测为正的样本数; “TN”表示负样本中预测为负的样本数; “FP”表示负样本中预测为正的样本数; “FN”表示正样本中预测为负的样本数; “Accuracy”越接近 1 表示模型分类效果越好。

3.2 数据集

实验所使用的数据集中共包含两类: 一类为近红外活体样本; 另一类为近红外非活体样本。其中, 活体图像数据为香港理工大学在 2010 年采集的公开数据集 PolyU-NIRFD^[10], 该数据集包含了 350 名志愿者分别在不同光照强度、不同头部姿态和不同表情下的近红外图像, 每人被采集了约 100 张图像, 共计 38981 张, 使用 MTCNN (Multi-task convolutional neural network) 人脸检测器^[11]截出人脸区域。相关示例样本如图 2(a)所示。

非活体样本图像数据为本文自建数据集, 该数据集建立方式遵循以下准则:

1) 图像攻击数据应具有多样性。本文从 CelebA 数据集、互联网中随机选取了约 500 张不同头部姿态和不同表情状态下的图像, 并使用惠普打印机 (DeskJet 2600) 将每张图像打印在 A4 纸上, 以构建攻击样本。

2) 采集过程应与活体样本采集过程类似。根据 PolyU-NIRFD 的采集条件, 本文在数据采集时控制了以下变量:

①距离特定。在和摄像头距离保持不变的情况下, 随机开启或关闭摄像头补光灯, 在这两种光照条件下, 手持 A4 纸图像, 通过旋转、弯曲纸张等方式进行图像捕捉。

②光照特定。光照包含两种情况, 一种是打开补光灯, 另一种是关闭补光灯, 在补光灯开启/关闭的情况下, 手持 A4 纸张由远及近、从左往右的移动, 在

移动过程中, 随机旋转、弯曲纸张。

上述采集过程设置每 5 帧保存一张图片, 最终利用近红外摄像头采集了 A4 纸图像在不同距离、不同角度、不同光照下的近红外图像, 共计 42300 张, 使用 MTCNN 人脸检测器截出人脸区域。相关示例样本如图 2(b)所示。

3.3 模型训练

为了更好地验证模型的识别率, 本文采用 10 折交叉验证法, 每份样本集中的图像均采用随机抽样方法, 模型训练时, 设置 BatchSize 为 64, 共训练 10 个 epoch, 初始学习率为 10^{-4} , 每隔 5 个 epoch 学习率乘以 0.1, 梯度下降优化器为 Adam^[12], 将三通道人脸图像设为 128×128 大小, 使用最小最大值归一化方法将像素值归一化到 [0,1] 区间, 具体公式为:

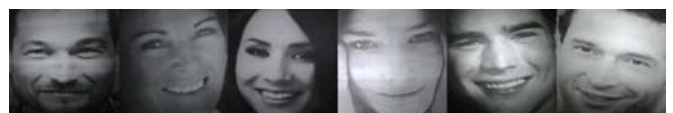
$$norm = \frac{x - x_{min}}{x_{max} - x_{min}} \quad (8)$$

式中: x 为当前像素值; x_{min} 、 x_{max} 分别为图像像素最小值和最大值。



(a) 近红外活体人脸图片样本

(a) Near-infrared liveness face image sample



(b) 近红外照片打印人脸图片样本

(b) NIR photo printed face picture sample

图 2 近红外活体检测数据示例

Fig.2 Examples of near-infrared liveness detection data

表 2 为 10 折交叉验证相关结果, 从中可得, 10 组测试集的准确率均在 99.90% 以上, 平均准确率达到 99.95%, 每组测试的准确率差距较小, 这说明模型在 epoch 达到 10 时已趋于稳定状态。图 3 为第一组实验的迭代过程图, 从图中可以看出, 当 epoch 大于 4 后, Accuracy 基本保持在一条直线上, Loss 值也下降到 0.02 以下, 随着迭代次数的增加, Loss 值继续下降。当 epoch 大于 5 后, Loss 值基本保持一条直线, 且十分接近于 0, 最终达到稳态。图 4 是测试图片过程中的所有卷积层提取的特征图可视化结果。

3.4 模型比较与结果分析

为了说明本文设计的 LeNet_Liveness 的有效性, 在数据预处理、数据集划分等基本条件不变的情况下, 本文分别使用 SVM 和 LeNet-5 对活体检测数据集进行分类识别, 最终实验测试结果如表 3 所示。

表 2 10 折交叉验证结果
Table 2 10-fold cross-validation results

Category	Test Dataset									
	1	2	3	4	5	6	7	8	9	10
Liveness	99.97	99.97	99.97	100	100	99.97	99.97	99.97	100	99.95
Non-liveness	99.98	99.95	99.91	99.98	99.88	99.93	99.86	99.93	99.91	99.83
Overall	99.96	99.95	99.94	99.95	99.98	99.95	99.91	99.95	99.96	99.90

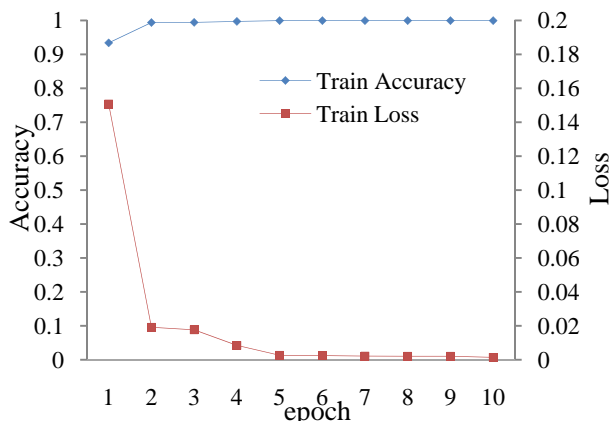


图 3 数据训练过程

Fig.3 Data training process

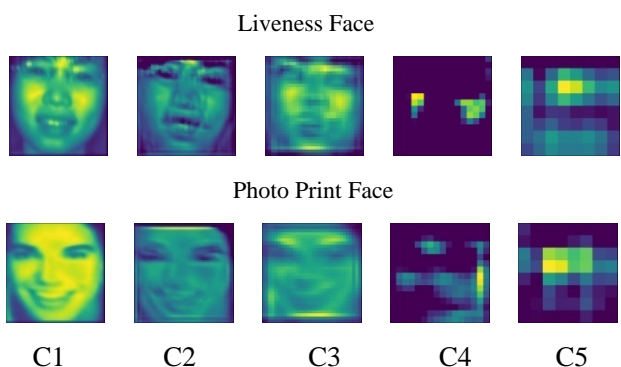


图 4 卷积层相关特征

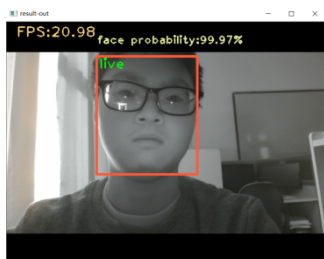
Fig.4 Convolution layer related features

由表 3 可得，本文的近红外数据集在 SVM 和 LeNet-5 的算法下均有较高的准确率，其中，SVM 的平均准确率为 96.67%，LeNet-5 的平均准确率为 98.23%，这是因为真人/照片在近红外光成像中特征明显，易于区分，所以准确率较高。本文所提出的 LeNet_Liveness 的实验结果最高，达到了 99.95%，单张图片的平均预测时间在 CPU 上为 31.08 ms，GPU 上仅为 10.77 ms，与 SVM 和 LeNet-5 相比，速度较慢，这是由于模型结构的加深、卷积核数量的增加等带来的弊端，降低了推理速度，但对于 FPS 为 30 的摄像头来说，实时性依旧较高。图 5 为本文的近红外活体检测系统，当真人人脸/照片打印出现在系统中时，系统会自动判断对象的活体属性，且系统的实时速度每秒约为 18~22 帧（包含了 MTCNN 人脸检测算法），实时性较高。

表 3 三种算法结果比较

Table 3 Comparison of the results of the three algorithms

Algorithm	Accuracy/%	Average prediction time for a single picture/ms	
		GPU	CPU
SVM	96.67	-	4.43
LeNet-5	98.23	2.03	7.57
LeNet_Liveness	99.95	10.77	31.08



(a) 真人识别结果

(a) Real person recognition result



(b) 照片攻击识别结果

(b) Photo attack recognition result

图 5 活体检测系统示例

Fig.5 Examples of live detection systems

注：系统测试对象均未在数据集中出现过（Note: None of the system test objects have appeared in the dataset）

为了进一步验证本文方法的有效性，现与其它文献中所提出的活体检测方法进行对比，对比结果如表4所示。

由表4各项指标比较可得，交互式活体检测基本不依赖于特殊设备，环境易部署，成本低，算法准确率高，但人机友好性较差。静默活体检测大多依赖于

特殊设备，特别是近红外设备，在近红外光下，使用CNN提取的特征较其它机器学习类算法明显，具有很高的活体识别率，本文借助近红外图像和CNN取得的活体识别率最高，达到了99.95%，具有良好的鲁棒性和泛化能力。

表4 不同文献结果比较
Table 4 Comparison of results from different literature

Detection type	Literature	Algorithm	Equipment	Accuracy/%
Interactive	[2]	Head posture + mouth opening and closing detection	Visible light camera	99.25
	[4]	Random emoji commands	Visible light camera	95.85
	[5]	Blink detection + smile detection + open mouth detection	VTM camera	97.67
Silent	[13]	LBP+Gabor+SVM	Visible light camera	98.00
	[14]	SVM+3D point cloud reconstruction+Face key point	Binocular camera(Near infrared light+visible light)	99.00
	[15]	CNN(double-mean pooling +multiple types of activation function)	Visible light camera	99.67
	This article	CNN (LeNet-5 improvements)	near-infrared camera(Near infrared light)	99.95

4 结语

根据近红外光在真人人脸和打印照片上的成像特点，提出一种基于改进 LeNet-5 和近红外图像的静默活体检测方法，在 LeNet-5 的结构基础上进行改进，构建了一个 LeNet_Liveness，针对该结构进行了多次的实验分析，最终实验发现，在近红外场景下，本文提出的 LeNet_Liveness 在活体检测数据集上有较高的分类准确率，对抗非活体攻击非常有效。

虽然本文对 LeNet-5 进行了结构上的优化，增加了网络训练、推理时间和模型参数，但在活体识别率上取得了比较好的成绩。高精度、高效率一直是科研人员在深度学习领域不断探索的最终目标，本文虽然提高了模型精度，但在推理效率上依旧有所降低，下一步，本文计划在维持模型识别率的基础上，减少模型参数，提升推理效率，同时，配合可见光摄像头，利用 CNN、多色彩特征等方法，实现在可见光场景下的静默活体检测。

参考文献

[1] Singh A K, Joshi P, Nandi G C. Face recognition with liveness detection using eye and mouth movement [C]//*Proceedings of the 2014 International*

Conference on Signal Propagation and Computer Technology (ICSPCT), IEEE, 2014: 592-597.
 [2] 张进, 张娜娜. 优化特征提取的交互式人脸活体检测研究[J]. *计算机工程与应用*, 2019, 55(13): 193-200.
 ZHANG Jin, ZHANG Nana. Research on Interactive Face Detection Based on Optimized Feature Extraction[J]. *Computer Engineering and Applications*, 2019, 55(13): 193-200.
 [3] V David, A Sanchez. Advanced support vector machines and kernel methods[J]. *Neurocomputing*, 2003, 55(1/2): 5-20.
 [4] Ng E S, Chia Y S. Face verification using temporal affective cues[C]//*Proceedings of the 21st International Conference on Pattern Recognition, Piscataway*, 2012: 1249-1252.
 [5] 马钰锡, 谭励, 董旭, 等. 面向 VTM 的交互式活体检测算法[J]. *计算机工程*, 2019, 45(3): 256-261.
 MAYuxi, TAN Li, DONG Xu, et al. Interactive Liveness Detection Algorithm for VTM[J]. *Computer Engineering*, 2019, 45(3): 256-261.
 [6] Lecun Y, Bottou L. Gradient-based learning applied to document recognition[C]//*Proceedings of the IEEE*, 1998, 86(11): 2278-2324.
 [7] 李文宽, 刘培玉, 朱振方, 等. 基于卷积神经网络和贝叶斯分类器的句子分类模型[J]. *计算机应用研究*, 2020, 37(2): 333-336, 341.
 LI Wenkuan, LIU Peiyu, ZHU Zhenfang, et al. Sentence classification model based on convolution neural network and Bayesian classifier[J].

- Application Research of Computers*, 2020, **37**(2): 333-336, 341.
- [8] 程淑红, 周斌. 基于改进 CNN 的铝轮毂背腔字符识别[J]. *计算机工程*, 2019, **45**(5): 182-186.
CHENG Shuhong, ZHOU Bing. Recognition of Characters in Aluminum Wheel Back Cavity Based on Improved Convolution Neural Network[J]. *Computer Engineering*, 2019, **45**(5): 182-186.
- [9] LIN M, CHEN Q, YAN S. Network In Network [EB/OL]. [2014-03-04]. <https://arxiv.org/pdf/1312.4400.pdf>.
- [10] ZHANG B, ZHANG L, ZHANG D, et al. Directional binary code with application to PolyU near-infrared face database[J]. *Pattern Recognition Letters*, 2010, **31**(14):2337-2344.
- [11] 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.
- [12] KINGMA D P, BA J. Adam: a method for stochastic optimization[EB/OL]. [2014-12-22]. <https://arxiv.org/pdf/1412.6980v8.pdf>.
- [13] Määttä J, Hadid A, Pietikäinen M. Face spoofing detection from single images using micro-texture analysis[C]//*Proceedings of the International Joint Conference on Biometrics, IEEE*, 2011: 1-7.
- [14] 邓茜文, 冯子亮, 邱晨鹏. 基于近红外与可见光双目视觉的活体人脸检测方法[J]. *计算机应用*, 2020, **40**(7): 2096-2103.
DENG Qianwen, FENG Ziliang, QIU Pengchen. Face liveness detection method based on near-infrared and visible binocular vision[J]. *Journal of Computer Applications*, 2020, **40**(7): 2096-2103.
- [15] 龙敏, 佟越洋. 应用卷积神经网络的人脸活体检测算法研究[J]. *计算机科学与探索*, 2018, **12**(10): 1658-1670.
LONG Min, TONG Yueyang. Research on Face Liveness Detection Algorithm Using Convolutional Neural Network[J]. *Journal of Frontiers of Computer Science and Technology*, 2018, **12**(10): 1658-1670.