计算机系统应用 ISSN 1003-3254, CODEN CSAOBN
Computer Systems & Applications,2018,27(8):35−41 [doi: 10.15888/j.cnki.csa.006491]

E-mail: csa@iscas.ac.cn
http://www.c-s-a.org.cn
Tel: +86-10-62661041

基于残差量化卷积神经网络的人脸识别方法①

周光朕[1], 杜姗姗[2], 冯　瑞[1,2], 欧丽君[3], 刘　斌[4]

[1](复旦大学 计算机科学技术学院, 上海 201203)
[2](上海市智能信息处理重点实验室 上海视频技术与系统工程研究中心, 上海 201203)
[3](上海临港智慧城市发展中心, 上海 201306)
[4](上海无线电设备研究所, 上海 200090)

摘　要: 针对大规模人脸识别问题, 基于残差学习的超深卷积神经网络模型能取得比其他方法更高的识别精度, 然而模型中存在的海量浮点参数需要占用大量的计算和存储资源, 无法满足资源受限的场合需求. 针对这一问题, 本文设计了一种基于网络参数量化的超深残差网络模型. 具体在 Face-ResNet 模型的基础上, 增加了批归一化层和 dropout 层, 加深了网络层次, 对网络模型参数进行了二值量化, 在模型识别精度损失极小的情况下, 大幅压缩了模型大小并提升了计算效率. 通过理论分析与实验验证了本文设计方法的有效性.

关键词: 人脸识别; 残差学习; 量化卷积神经网络

引用格式: 周光朕,杜姗姗,冯瑞,欧丽君,刘斌.基于残差量化卷积神经网络的人脸识别方法.计算机系统应用,2018,27(8):35–41. http://www.c-s-a.org.cn/1003-3254/6491.html

Face Recognition Method Based on Quantized Residual Convolutional Neural Networks

ZHOU Guang-Zhen[1], DU Shan-Shan[2], FENG Rui[1,2], OU Li-Jun[3], LIU Bin[4]

[1](School of Computer Science, Fudan University, Shanghai 201203, China)
[2](Shanghai Key Laboratory of Intelligent Information Processing, Shanghai Engineering Research Center for Video Technology and System, Shanghai 201203, China)
[3](Shanghai Lingang Smart City Development Center, Shanghai 201306, China)
[4](Shanghai Radio Equipment Research Institute, Shanghai 200090, China)

Abstract: Very deep convolutional neural networks based on residual learning have achieved higher accuracy than other methods for large scale face recognition problem. But the massive floating-point parameters existing in the models need to occupy extensive computational and memory resources, which cannot be satisfied with the demand of occasions with limited resources. Aimed at the solution of this issue, a very deep residual neural network based on network model parameters quantization was designed in this study. In detail, based on the model Face-ResNet, the network was added with batch normalization layers and dropout layers, and also its total layers were deepened. Applying binary quantization to parameters of the designed network models, it can compress the model size substantially and improve computational efficiency with little loss of model recognition accuracy. Both theoretical analysis and experiments prove the effectiveness of the designed method.

Key words: face recognition; residual learning; quantized convolutional neural networks

① 基金项目: 国家重点研发计划(2017YFC0803700); 上海市科委项目(17511101702); 临港地区智能制造产业专项(#ZN2016020103); 复旦大学工程与应用技术研究院先导项目 (gyy2017-003)

Foundation item: National Key Research and Development Project of China (2017YFC0803700); Program of Shanghai Committee of Science and Technology (17511101702); Special Projct in Intelligent Manufacturing of Lingang Region (#ZN2016020103); Guided Project of Academy for Engineering and Technology, Fudan University (gyy2017-003)

收稿时间: 2018-01-02; 修改时间: 2018-01-23; 采用时间: 2018-01-25; csa 在线出版时间: 2018-07-28

1 引言

人脸识别是一种重要的生物识别技术. 利用人脸特征信息, 可实现用户无感的身份识别, 具有广阔应用前景. 过去数十年间, 许多人脸识别方法被提出, 早期方法通常利用人工设计的特征, 结合不同分类器和度量方法进行识别, 但识别精度并不理想, 存在较大提升空间.

近年来, 基于神经网络的方法被广泛用于人脸识别[1,2], 取得较好的识别精度. 文献[3]提出基于残差学习的卷积神经网络 (ResNet), 这是一种具有良好学习与泛化能力的网络模型, 也可被用于人脸识别[4]. 文献[4]采用残差卷积神经网络结构, 用三个公开数据集总计 17 189 人, 约 70 万张图像在交叉熵损失函数监督下训练, 在人脸识别数据集 LFW[5]上取得优异的识别精度.

然而残差卷积神经网络模型中存在海量浮点参数, 应用时要消耗巨大的计算和存储资源, 效率低下. 例如 ResNet-50 有超过 2500 万的参数, 对 224×224 大小的 3 通道输入, 计算总量需要 42 亿多次的浮点乘法. 对网络模型进行参数量化是一类可以有效降低网络模型存储消耗和提升计算效率的方法[6], 包括二值神经网络 BNN[7]、异或网络 XNOR-Net[8]和多位量化网络 DoReFa-Net[9]等. BNN 根据模型参数的符号, 用±1 表示, 并以符号变换甚至异或替代卷积的浮点乘法, 大幅压缩了网络模型并提升了计算速度, 但有较大的模型精度损失. XNOR-Net 在 BNN 基础上对每层引入一个浮点系数, 来减少量化导致的模型精度损失. DoReFa-Net 则采用量化到多个比特位的方法, 能进一步减少模型精度损失, 但相比前二者, 需要损失一定的计算加速和网络压缩比.

针对应用于大规模人脸识别的超深残差卷积神经网络存在网络模型过大和计算效率低下的问题, 设计了一种基于网络参数量化的超深残差网络模型, 包含残差网络模型选择和网络参数量化两部分. 具体在模型 Fase-ResNet (https://github.com/ydwen/caffe-face) 中添加批归一化层和 dropout 层, 并去除中心损失函数, 作为基准模型, 再通过残差构件堆加和内部层数加深提升基准模型识别精度, 并对精度最优的网络参数进行二值量化, 在模型识别精度损失极小的情况下, 大幅压缩了网络模型大小, 并提升了计算效率.

本文组织如下: 在第 2 节介绍了本文工作的算法设计框架; 在第 3 节和第 4 节分别介绍了框架中两大部分工作, 残差卷积神经网络模型选择和卷积神经网络参数量化; 在第 5 节通过理论分析和实验研究分析了本文所设计网络模型的精度、相对基准模型的网络压缩比和效率提升比; 最后在第 6 节对全文工作进行了总结.

2 算法设计框架

基于网络参数量化的超深残差网络模型的设计包括两部分内容: 残差卷积神经网络的模型选择和卷积神经网络的参数量化如图 1 所示. 图 1 的上部和实线右侧图 1(c) 到 (f) 给出了其主要工作流程及网络结构, 实线左侧图 1(a) 和 (b) 是结构 (c) 到 (f) 中使用的不同残差构件的内部细节.

模型选择部分对应图 1 上部左侧三个方框的内容. 在 Fase-ResNet 的训练模型中加上批归一化 (BN) 层和 dropout 层以加快训练收敛和防止过拟合, 并去除中心损失函数, 以此作为基准模型 A, 对应图 1(c) 的网络结构. 结合残差网络结构加深的方法, 依次在 B 和 C 模型中加深了残差构件的堆积次数和内部层数, 分别对应图 1(d) 和 (e) 的网络结构. 其中 A 和 B 网络采用图 1(a) 的残差构件, C 网络采用图 1(b) 的残差构件. 在图 1 中, 标识了 B 和 C 网络与 A 使用不同数量和类型残差构件的区别, 三者的 dropout 层应用于 Fc5 层.

网络量化部分对应图 1 上部最右侧方框内容, 采用了基于 BNN、XNOR-Net 和 DoReFa-Net 中的量化函数对神经网络模型的参数进行了量化, 并参考其实验方法训练量化网络. 考虑到量化操作会一定程度降低原网络模型的识别精度, 保留了网络第一个卷积层 Conv1a 和最后一个全连接层 Fc6 的浮点数参数, 其他量化层在图 1(f) 中加上了 Q 作为标记. 通过对模型选择工作设计的三个网络模型识别精度最优者进行了网络参数的二值量化, 实现在极小的识别精度损失情况下, 大幅压缩网络模型大小并提升计算效率.

3 残差卷积神经网络模型选择

3.1 残差学习机制

残差学习在残差神经网络 ResNet 中被提出, 是一种简单高效的网络模型学习机制. 深度卷积神经网络通常直接学习输入数据到输出标签的目标映射, 但简单地加深这类网络却可能出现训练精度不升反降的情况, 即网络模型的恶化 (degradation) 问题, 且该问题并

非由过拟合引起. 残差学习机制则学习目标映射与原输入的残差量, 通过残差值与原输入相加恢复最终的目标映射, 可有效解决这一恶化问题.

令$H(x)$表示输入x的目标映射, 卷积神经网络学习的目标函数为$F(x)$. 常见方法直接学习这个目标映射, 即$F(x)=H(x)$. 当恒等映射$H(x)=x$为目标映射时, $F(x)$较难拟合这个目标. 残差学习机制则学习目标映射与输入之间的残差量, 即$F(x)=H(x)-x$, 而目标映射则可以简单地通过原输入与残差量相加来恢复, 即$H(x)=F(x)+x$.

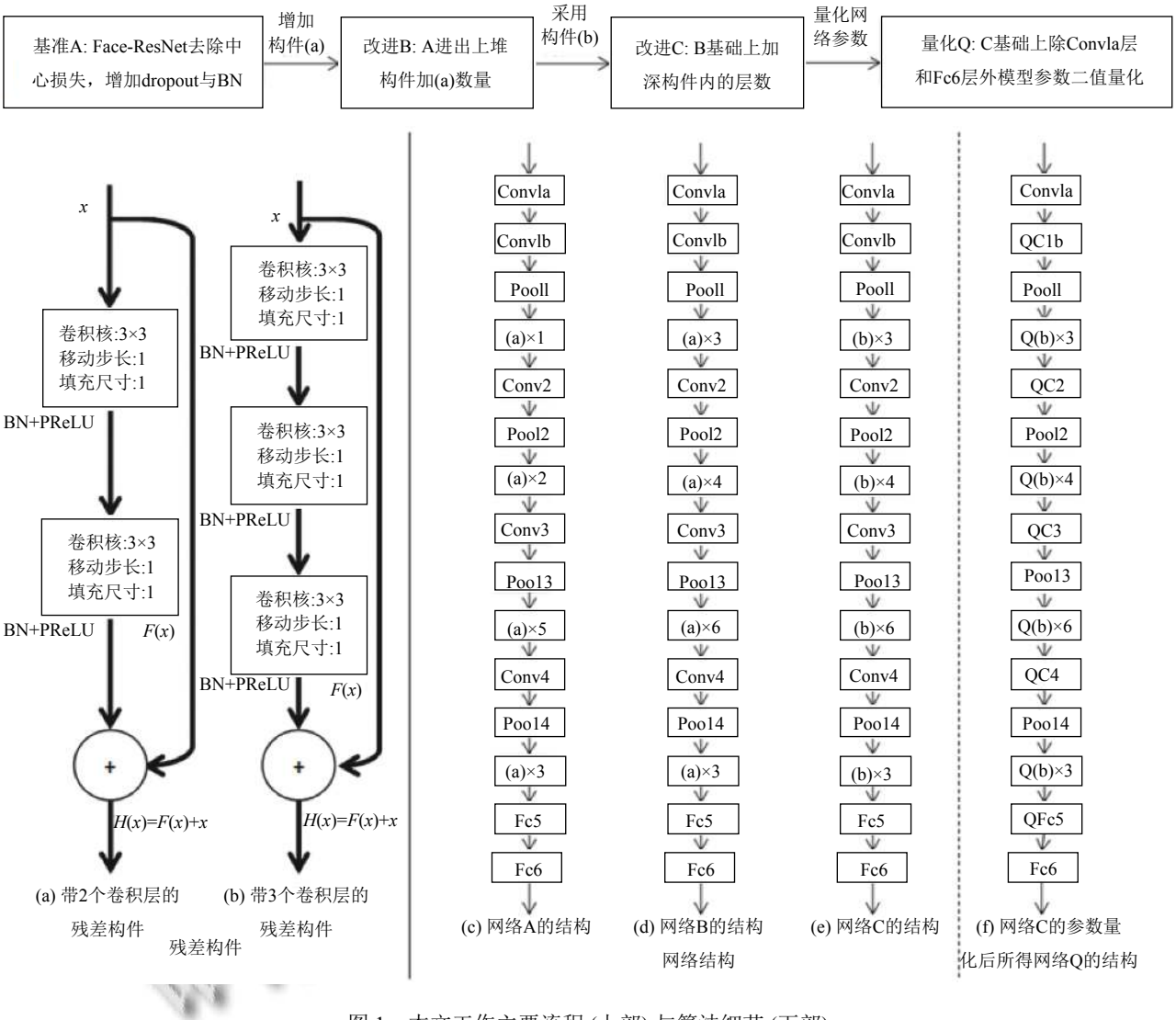

图 1 本文工作主要流程 (上部) 与算法细节 (下部)

完成如上残差量与原输入相加操作的一个模块被称为残差构件. 设计不同的残差量学习函数$F(x)$可得到不同形式的构件, 图 1 中 (a) 和 (b) 就是可能的两种. 图中的纵向卷积分支用于学习残差量$F(x)$, 而原输入通过旁路分支直接与残差量相加, 来恢复目标映射$H(x)$. 同时, 在每个卷积层的激活函数前添加批归一化层可加快网络模型的训练速度.

残差学习机制在 ResNet 中被证明可用来训练超深层次的卷积神经网络, 精度优于其他常见的卷积神经网络方法, 且网络模型层次越深, 精度越好.

3.2 模型选择过程

针对 Face-ResNet 网络设计基准模型存在的两个问题: 一是此模型的全连接层存在较多参数, dropout 层被引入来防止模型发生过拟合情况; 二是为加快网络训练, 在每个卷积层的激活函数前添加批归一化层.

但是, 网络中原有的中心损失函数需要计算全连接层输出特征与原数据空间每一类中心值之间的差异, 而 Dropout 层却按设定比例随机舍弃全连接层参数, 将导致训练时全连接层输出特征分布非常不稳定, 极大影响中心损失函数的监督效果. 另外, 中心损失函数需要得到训练数据每类中心. 若存在离群值, 将对中心的位置造成影响, 进而影响中心损失函数的监督效果. 基于这两个问题, 本文在基准模型中舍去中心损失函数, 只用交叉熵损失函数作为监督. 经过以上改进可得网络模型 A, 图 1(a) 是 A 使用的残差构件, 图 1(c) 是其网络大致结构.

根据残差卷积神经网络层次越深, 精度越好的特点, 本文在网络 A 基础上加深网络模型. 但过深的残差网会参数过量, 训练数据量一定时, 网络可能得不到充分训练而使精度的提升并不多, 同时却要消耗更多存储和计算资源, 影响量化带来的压缩比和效率提升, 故本文只参考了 34 层与 50 层 ResNet 的设置. 首先对 A 网络增加残差构件的堆积数, 在 Conv2_x、Conv3_x、Conv4_x 分别增加了 2 个、2 个、1 个图 1(a) 的残差构件, 得到网络 B. 接着, 又在 B 的基础上增加构件内部的卷积层数, 即用图 1(b) 的残差构件, 得到网络 C. 网络 B 和 C 的大致结构如图 1(d) 和 (e) 所示. 除以上设计, Face-ResNet 其他层的参数设均被保留.

残差构件中卷积层的卷积核、移动步长和填充尺寸分别是 3×3、1 和 1, 其他卷积层不作填充. 卷积支路的填充操作可保证原输入x和残差量$F(x)$相加时维度对应相等. 池化层的区间设为 2×2, 移动步长为 2. 在相同通道数的卷积层中间, 输出的特征图大小不变; 而特征图经过池化长宽均缩小一半时, 卷积层通道数则加倍. 网络模型 A 到 C 每层具体的参数设置见表 1, 其中输入均为 112×96 像素大小的 3 通道图片数据.

表 1 模型选择 3 种结构的具体参数设置

layer name	A	B	C
Input		112×96×3	
Conv1a		3×3, 通道 32, 步长为 1	
Conv1b		3×3, 通道 64, 步长为 1	
Pool1		2×2, 最大池化, 步长为 2	
Conv2_x	$\begin{bmatrix} 3\times3, \text{通道}64 \\ 3\times3, \text{通道}64 \end{bmatrix}\times1$	$\begin{bmatrix} 3\times3, \text{通道}64 \\ 3\times3, \text{通道}64 \end{bmatrix}\times3$	$\begin{bmatrix} 3\times3, \text{通道}64 \\ 3\times3, \text{通道}64 \\ 3\times3, \text{通道}64 \end{bmatrix}\times3$
Conv2		3×3, 通道 128, 步长为 1	
Pool2		2×2, 最大池化, 步长为 2	
Conv3_x	$\begin{bmatrix} 3\times3, \text{通道}128 \\ 3\times3, \text{通道}128 \end{bmatrix}\times2$	$\begin{bmatrix} 3\times3, \text{通道}128 \\ 3\times3, \text{通道}128 \end{bmatrix}\times4$	$\begin{bmatrix} 3\times3, \text{通道}128 \\ 3\times3, \text{通道}128 \\ 3\times3, \text{通道}128 \end{bmatrix}\times4$
Conv3		3×3, 通道 256, 步长为 1	
Pool3		2×2, 最大池化, 步长为 2	
Conv4_x	$\begin{bmatrix} 3\times3, \text{通道}256 \\ 3\times3, \text{通道}256 \end{bmatrix}\times5$	$\begin{bmatrix} 3\times3, \text{通道}256 \\ 3\times3, \text{通道}256 \end{bmatrix}\times6$	$\begin{bmatrix} 3\times3, \text{通道}256 \\ 3\times3, \text{通道}256 \\ 3\times3, \text{通道}256 \end{bmatrix}\times6$
Conv4		3×3, 通道 512, 步长为 1	
Pool4		2×2, 最大池化, 步长为 2	
Conv5_x	$\begin{bmatrix} 3\times3, \text{通道}512 \\ 3\times3, \text{通道}512 \end{bmatrix}\times3$	$\begin{bmatrix} 3\times3, \text{通道}512 \\ 3\times3, \text{通道}512 \end{bmatrix}\times3$	$\begin{bmatrix} 3\times3, \text{通道}512 \\ 3\times3, \text{通道}512 \\ 3\times3, \text{通道}512 \end{bmatrix}\times3$
Fc5		1024	
Dropout		比例 0.5	
Fc6		10575	
Softmax		-	

4 卷积神经网络参数量化

卷积神经网络参数量化方法通常将网络每层参数量化到 k 个值, 用$\lceil \log_2 k \rceil$位索引表示, 量化模型保存索引与这些值. 取整和符号函数常用于量化. 若 k=2, 且取值为±1, 则为二值量化. 二值量化网络只用 1 比特位表示参数, 可大幅压缩网络模型, 计算时用符号改变替代浮

点乘法, 可提高模型的计算效率.

4.1 二值神经网络

假设卷积神经网络模型有 L 层, 记第$l(l=1,2,\cdots,L)$层权值参数矩阵为$W_l \in R^{k_l \times k_l \times c_l}$, 每个元素为$w$, 其中$k_l$为卷积核边长, c_l为通道数. 用W_l^b表示二值量化后网络每层参数, 每个元素$w^b \in \{+1,-1\}$. 二值神经网络 BNN 中使用了符号函数进行量化, 如公式 (1).

$$w^b = sign(w) = \begin{cases} +1, & \text{如果} w \geq 0 \\ -1, & \text{其他情况} \end{cases} \tag{1}$$

符号函数的导函数几乎处处为 0, 反向传播算法不能训练网络, 需要设定其梯度反传函数. BNN 在训练时保留绝对值较小参数的梯度而抑制绝对值较大参数的梯度, 以g_s和g_w分别表示符号函数和恒等函数的梯度, BNN 使用$g_s = g_w 1_{|w|\leqslant 1}$, 具体如公式 (2) 所示.

$$g(w)1_{|w|\leqslant 1} = \begin{cases} 0, & w > 1 \\ 1, & w \in [-1,1] \\ 0, & w < -1 \end{cases} \tag{2}$$

将以上两者用于普通卷积神经网络训练中, BNN 对网络每层参数进行二值量化, 可一定程度保持原浮点网络模型的精度, 同时大幅压缩网络模型并提升计算效率.

4.2 异或网络

由于 BNN 应用于大规模数据集上将造成网络模型精度大幅下降, 异或网络 XNOR-Net 做了改进. 离散卷积操作可转为矩阵相乘. 令每层的输入矩阵为$\mathbf{X}$, 参数矩阵为$\mathbf{W}$, 则每层卷积计算为$\mathbf{W}\cdot\mathbf{X}$. 异或网络在每层二值矩阵$\mathbf{W}^b$前乘上一个浮点系数$\alpha$, 卷积计算近似变为$\mathbf{W}\cdot\mathbf{X} \approx \alpha(\mathbf{W}^b\cdot\mathbf{X})$. 选取最佳$\alpha$和$\mathbf{W}^b$近似原卷积计算则通过固定$\mathbf{W}$后最小化$J(\mathbf{W}^b,\alpha) = \left\|\mathbf{W}-\alpha\mathbf{W}^b\right\|^2$得到. 利用$\mathbf{W}^{b\mathrm{T}}\cdot\mathbf{W}^b = N$($N$ 为卷积核元素个数), 可得到最优解$\mathbf{W}^{b*} = sign(\mathbf{W})$和$\alpha^* = \|W\|/N$. $\mathbf{W}^b$仍由$\mathbf{W}$用符号函数二值量化, α为每层卷积矩阵参数绝对值和的平均.

神经网络中间层的输入为上一层输出的激活值. 按同样方式, 可得到每层输入$\mathbf{X}$的二值量化近似表示$\mathbf{X} \approx \beta\mathbf{X}^b$. 此时, 原来的卷积计算就可以近似变为$\mathbf{W}\cdot\mathbf{X} \approx \alpha\beta(\mathbf{W}^b\cdot\mathbf{X}^b)$. 其中$\mathbf{W}^b\cdot\mathbf{X}^b$可用异或操作计算, 计算速度大幅提升.

卷积神经网络的卷积模块包含四部分, 依次为卷积、批归一化、激活和池化. 若按此顺序对神经网络参数和激活值二值量化, 且采用最大池化方式, 将导致多数+1 被保留而–1 被舍去, 对模型精度会造成巨大影响. 因此, 将池化放在激活之前, 改动卷积模块为卷积、批归一化、池化和激活, 再对卷积参数和池化输出作二值量化. 此外, 将输入归一化, 使其均值为零, 也可减少二值量化带来的模型精度损失.

异或网络可用 BNN 的训练方法, 给每层二值量化参数和激活值乘上一个浮点系数, 相比 BNN, 可在网络模型精度更小损失的情况下获得相近的网络压缩比和计算效率提升.

4.3 多位量化网络

在前两种量化网络基础上, 多位量化网络 DoReFa-Net 提出对网络参数进行多位量化的方法. 设$x \in R$是闭区间[0, 1]上的一个实数, 将它量化到 0 到 1 闭区间上的k 位 ($k \geqslant 2$) 量化值x^q的函数$q_k(x)$如公式 (3).

$$x^q = q_k(x) = \frac{1}{2^k-1} round((2^k-1)x) \tag{3}$$

神经网络每层参数W和激活值A的量化函数如公式 (4) 和 (5) 所示.

$$W^q = 2q_k\left(\frac{\tanh(W)}{2\max(|\tanh(W)|)} + \frac{1}{2}\right) - 1 \tag{4}$$

$$A^q = q_k(h(A)) \tag{5}$$

其中, $h(x)$是值域在[0, 1]上的函数, 可用的有 $h(x) = 0.5(\tanh(x)+1)$, $h(x) = \min(1,|x|)$, $h(x) = clip(x,0,1)$ 等. 当 k=1 时, 则使用异或网络的二值量化方法.

DoReFa-Net 使用前两种量化网络的训练方式, 设置网络参数的不同量化位数 k, 可对网络模型进行不同程度的量化. 相比 BNN 和 XNOR-Net, 量化位数 k 越大, 对原浮点网络模型精度导致的损失越小, 但同时网络模型也越大, 需要更多存储空间, 对计算效率的提升也越少. 需要设置合理的 k 值折衷网络精度损失和网络模型压缩比与计算效率提升.

DoReFa-Net 也有对训练梯度的多位量化方法, 但本文考虑到梯度对训练的重要性, 保留其浮点表示. 本文基于多位量化网络的参数量化方法, 在实验中对网络模型参数进行二值量化. 实验情况将在第 5 节中说明.

5 实验与分析

5.1 实验环境与数据集准备

本文算法实验的系统环境为 Linux Ubuntu 16.04, 服务器配置为 Intel Xeon E5-2620 v4 处理器、128 G 内存

和 NVIDIA 图形处理单元 GTX 1080 TI GPU 卡. 实验采用 tensorpack(https://github.com/ppwwyyxx/tensorpack)工具包搭建本文设计的网络模型.

CASIA-WebFace 数据集[10]被用于实验的训练, 利用室外带标注的人脸数据集 LFW 用于网络模型的精度测试. CASIA-WebFace 是目前最大的公开人脸识别数据集之一, 全部数据采集自互联网, 包含 10 575 人共 494 414 张图片. LFW 数据集则有 5749 位名人的 13 233 张网络图片, 均采集自室外场景.

一些方法被用于训练和测试数据的预处理. 对 CASIA-WebFace 的标注 (http://zhengyingbin.cc/ActiveAnnotationLearning/)[11]被用于提升训练数据质量. MTCNN(https://github.com/kpzhang93/MTCNN_face_detection_alignment)[12]则被用于检测图像中人脸位置和人脸特征点, 如双眼中心、嘴角两侧等. 文献[13]方法 (https://github.com/AlfredXiangWu/face_verification_experiment) 可用特征点进行对齐, 将人脸转为基本正视前方. 若 MTCNN 未检测到人脸, 直接舍弃训练集图片, 而对测试集图片截取中心 144×144 像素区域并对齐. 训练数据集经过处理后余下约 44 万张图片, 图像被调整到 128×128 大小, 测试集图片被调整到 144×144 大小.

图像输入网络时被裁剪到 112×96 大小, 训练集使用随机裁剪, 测试集采用中心裁剪. 另外, 实验对训练和测试数据的每个像素减去 127.5 并除以 128, 将像素归一化到–1 与+1 之间, 使其整体上具有均值为零的分布, 尽可能减少量化带来的网络精度损失.

5.2 实验设置与评判方式

随机抽取训练数据集每人两张图片, 共 21 150 张图片作为训练的验证集, 剩余图片用于训练. 训练中, 除使用随机裁剪外, 输入图像的翻转被用于数据增强. 网络模型训练阶段共 100 个 epoch, 每个 epoch 有 10 000 个 step, 每个 step 输入一个批次图像. 实验设置每批次 64 张图, 初始学习率为 0.01, 在 40、70、90 个 epoch 时将学习率依次缩小 10 倍.

量化网络模型的训练采用相同设置, 并采用多位量化网络的量化设置, 对 Conv1a 层和 Fc6 层外的卷积和全连接层进行二值量化. 实验中, 未量化网络约在 50 个 epoch 后收敛, 而量化网络则在约 55 个 epoch 后收敛.

由于两数据集的人并不相同, 测试时先对测试图片和其翻转输入网络模型, 取 Fc5 层的输出, 得到两组 1024 维特征, 拼接得到 2048 维特征. 判断两张测试图片是否为同一人时, 对两组 2048 维特征计算余弦距离相似度, 按照设定阈值进行判断.

LFW 提供一份分为 10 组随机生成的共 6000 对图像的测试集合验证算法精度. 其中 3000 对是相同的人, 另 3000 对是不同的人. 通过以上方法对 6000 对图像进行测试, 算法在 10 个分组中交叉验证, 以 10 组验证结果的均值报告算法精度.

5.3 实验结果

首先对模型选择中设计的三个网络模型进行了精度测评, 结果见表 2. 可以看到, 随着网络层次的增加, 基于残差学习的超深卷积神经网络模型的精度可以逐步提升, 并且加深后的网络 B 和 C 都超过了其他两个神经网络方法. 其中网络 C 精度达到 97.83%, 超过 LFW 报告的人类水平 97.53%, 是这些网络模型中的精度最优者.

接着, 对精度最优的网络 C 根据 5.2 节的设置进行了量化训练, 得到了量化模型 Q, 并在测试集上报告了 97.43% 的准确率, 相比模型 C 仅有 0.4% 的模型精度损失, 而比基准网络模型 A 更有精度的提升.

表 2 各种方法在 LFW 验证集上报告的精度结果

方法	准确率
DeepFace single[1]	0.9592
DeepFace ensemble[1]	0.9735
DeepID[2]	0.9745
Model A	0.9702
Model B	0.9747
Model C	**0.9783**
Model Q	0.9743
Human[5]	0.9753

5.4 压缩比与效率提升分析

由于测试时最后一层 Fc6 未参与计算, 故本文中针对模型计算效率提升的比较不含 Fc6 层. 参考异或网络文献, 量化网络 Q 相比网络 C 可有 2 倍的计算加速, 并能将网络压缩 32 倍. 表 3 列出了本文工作设计的 4 种网络结构的参数量、浮点计算量和相对基准模型 A 的压缩比与效率提升情况, 均通过理论计算得出. 可以看到, 尽管网络 C 相对 A 增加约 30% 的参数, 但经过二值量化, 模型 Q 相对 A 有整体上接近 4 倍的压缩. 当除去最后的 Fc6 层时, 压缩比甚至可以达到 22.8, 并有 1.05 倍的效率提升.

表 3 各网络模型的参数与相比基准模型的压缩比和效率提升情况

网络	参数量 ($\times10^7$)	浮点计算量 ($\times10^9$)	效率提升	压缩比 (整体)	压缩比
A	4.36	2.21	1	1	1
B	4.55	3.06	0.72	0.96	0.94
C	5.68	4.24	0.52	0.77	0.71
Q	5.68	2.12	1.05	3.81	22.8

6 总结

本文针对应用于大规模人脸识别的超深残差卷积神经网络存在网络模型过大和计算效率低下的问题, 设计了一种基于网络参数量化的超深残差网络模型, 具体在 Face-ResNet 模型基础上, 增加批归一化层和 dropout 层, 加深网络层次, 对网络模型参数进行二值量化. 通过理论分析和实验验证, 本文设计的网络可以在模型识别精度损失极小的情况下, 将网络模型压缩 22.8 倍并提升 1.05 倍计算效率. 由于目前仅量化了网络的参数, 未来可以考虑调整训练方式, 将计算时每层输入也进行量化, 以提升更多的计算效率. 量化的网络也适用于 FPGA, 可以考虑将其移植到 FPGA 上来进一步提升计算效率.

参考文献

1 Taigman Y, Yang M, Ranzato M, *et al*. DeepFace: Closing the gap to human-level performance in face verification. Proceedings of 2014 Conference on Computer Vision and Pattern Recognition. Columbus, OH, USA. 2014. 1701–1708.

2 Sun Y, Wang XG, Tang XO. Deep learning face representation from predicting 10,000 classes. Proceedings of 2014 IEEE Conference on Computer Vision and Pattern Recognition. Columbus, OH, USA. 2014. 1891–1898.

3 He KM, Zhang XY, Ren SQ, *et al*. Deep residual learning for image recognition. Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, NV, USA. 2016. 770–778.

4 Wen YD, Zhang KP, Li ZF, *et al*. A discriminative feature learning approach for deep face recognition. Proceedings of the 14th European Conference on Computer Vision. Amsterdam, The Netherlands. 2016. 499–515.

5 Huang GB, Mattar M, Berg T, *et al*. Labeled faces in the wild: A database forstudying face recognition in unconstrained environments. Proceedings of Workshop on Faces in 'Real-Life' Images: Detection, Alignment, and Recognition. Marseille, France. 2008.

6 雷杰, 高鑫, 宋杰, 等. 深度网络模型压缩综述. 软件学报, 2018, 29(2): 251–266. [doi: 10.13328/j.cnki.jos.005428]

7 Courbariaux M, Hubara I, Soudry D, *et al*. Binarized neural networks: Training deep neural networks with weights and activations constrained to +1 or -1. arXiv: 1602.02830, 2016.

8 Rastegari M, Ordonez V, Redmon J, *et al*. Xnor-net: Imagenet classification using binary convolutional neural networks. Proceedings of the 14th European Conference on Computer Vision. Amsterdam, The Netherlands. 2016. 525–542.

9 Zhou SC, Wu YX, Ni ZK, *et al*. DoReFa-Net: Training low bitwidth convolutional neural networks with low bitwidth gradients. arXiv: 1606.06160, 2016.

10 Yi D, Lei Z, Liao SC, *et al*. Learning face representation from scratch. arXiv: 1411.7923, 2014.

11 Ye H, Shao WY, Wang H, *et al*. Face recognition via active annotation and learning. Proceedings of 2016 ACM on Multimedia Conference. Amsterdam, The Netherlands. 2016. 1058–1062.

12 Zhang KP, Zhang ZP, Li ZF, *et al*. Joint face detection and alignment using multitask cascaded convolutional networks. IEEE Signal Processing Letters, 2016, 23(10): 1499–1503. [doi: 10.1109/LSP.2016.2603342]

13 Wu X, He R, Sun ZN. A lightened CNN for deep face representation. Proceedings of 2015 IEEE Conference on IEEE Computer Vision and Pattern Recognition. 2015.