

DOI:10.19651/j.cnki.emt.2209301

基于忆阻器卷积神经网络的表情识别*

赵益波^{1,2} 蒋文^{1,2} 孟若禹^{1,2} 李业宁^{1,2}

(1.南京信息工程大学电子与信息工程学院 南京 210044; 2.江苏省大气环境与装备技术协同创新中心 南京 210044)

摘要: 忆阻器具有纳米级尺寸、低功耗、类似神经突触等优点,在神经计算、图像分类等领域具有广阔的应用前景。本文提出了一种基于忆阻器卷积神经网络的面部表情识别方法,首先基于忆阻器构建了 ResNct 卷积神经网络,并对 ResNet 网络进行剪枝操作,然后将 ResNct 模型的所有卷积层以及全连接层的权重映射为忆阻器十字交叉阵列中忆阻器的忆导值。实验结果显示忆阻器卷积神经网络模型在 FER2013 数据集上的识别准确率为 63.82%,在 CK+数据集上的识别准确率为 93.95%。相比与原卷积网路,准确率损失仅分别为 0.31%和 0.76%。最后测试了忆阻器的非理想特性对准确率的影响,为忆阻器神经网络的实际部署提供参考。

关键词: 忆阻器神经网络;表情识别;ResNet;卷积神经网络

中图分类号: TP391.41 **文献标识码:** A **国家标准学科分类代码:** 510.40

Facial expression recognition based on memristor convolutional neural network

Zhao Yibo^{1,2} Jiang Wen^{1,2} Meng Ruoyu^{1,2} Li Yening^{1,2}

(1. School of Electronic & Information Engineering, Nanjing University of Information Science & Technology, Nanjing 210044, China; 2. Jiangsu Collaborative Innovation Center on Atmospheric Environment and Equipment, Nanjing 210044, China)

Abstract: Memristors have the advantages of nanoscale size, low power consumption and similar to neural synapses, etc., and have broad application prospects in neural computing, image classification and other fields. In this paper, a facial expression recognition method based on memristor-based convolutional neural network is proposed. First, a memristor-based ResNet convolutional neural network is constructed and the ResNet network is pruned. Then the weights of all convolutional layers and fully connected layers of the ResNet model are mapped as the memductance values of memristors in the memristive crisscross array. The experimental results show that the recognition accuracy of the memristor-based convolutional neural network model on the FER2013 dataset is 63.82%, and the recognition accuracy on the CK+dataset is 93.95%. Compared with the original convolutional network, the accuracy loss is only 0.31% and 0.76% respectively. Finally, the influence of the non-ideal characteristics of the memristor on the accuracy is tested, which provides a reference for the actual deployment of the memristor-based neural network.

Keywords: memristor neural network; facial expression recognition; ResNet; convolutional neural network

0 引言

当前,随着计算机计算能力的不断进步和数据集的丰富,深度学习在电子商务、制造业、农业和智能家居等各个领域得到了巨大发展^[1]。基于深度学习的智能应用程序通常依赖于强大计算能力的云数据中心。随着物联网设备和移动终端的不断推广,如何在资源受限的终端设备上部署高性能的深度学习模型已经引起了学术界和工业界的极大关注^[2]。边缘智能的概念由此应运而生^[3],在边缘环境中,受

物理尺寸和能源供应的限制,设备资源通常非常有限。因此,探索一种新的具有更好的记忆和处理性能的计算体系结构,建立有效的计算模型,是边缘智能发展的关键问题^[4]。

人脑的工作模式意味着一种非冯诺依曼计算架构,称为神经形态计算。神经形态计算可以从多个来源积累少量信息,产生不同类型的信号,并在需要时发出电流,模拟神经元的运行过程。神经形态计算将是高性能计算的下一个发展阶段,它可以极大地提高数据处理和机器学习的能力。近年来,新发现的忆阻器可以模拟大脑神经突触,是神经形

收稿日期:2022-03-16

* 基金项目:江苏高校优势学科 III 期建设工程项目、国家自然科学基金(61871230)、江苏省自然科学基金(BK20181410)项目资助

态计算元器件的上佳选择。忆阻器具有无源记忆功能,可以通过控制输入的磁通量或电荷来改变其阻值,功耗较低,使用纳米尺寸的忆阻器可以实现高性能的深度学习模型^[5]。忆阻器神经网络具有快速的信息存储和处理能力。2015年,加利福尼亚大学圣塔芭芭拉分校使用 12×12 的忆阻器交叉结构完成了对 3 个不同字母的分类^[6]。2017年,清华大学的研究小组使用了 128×8 的忆阻器十字交叉阵列来进行 3 张人脸的分类,准确率可大于 95%^[7]。2020年,该研究小组发表了一篇关于基于忆阻器十字交叉阵列的存储计算集成系统的研究成果^[8]。随着忆阻器神经网络的发展,将复杂深度学习模型通过忆阻器部署在边缘设备上逐渐成为可能。

在日常交际过程中,面部表情是个人表达自己内心情感最直接的方式,通过识别面部表情可以获得个人当时的身体或精神状态^[9]。因此,如何高效地识别面部表情,在人机交互、自动驾驶等领域具有重要意义。表情识别在深度学习当中可以归类为一个传统的分类问题^[10]。早期的网络架构 AlexNet^[11] 和 VGGNet^[12] 通过堆叠卷积运算符,但是过度地堆叠网络层数会出现梯度消失以及梯度爆炸等问题。GoogleNet^[13] 通过使用不同内核大小的卷积,增加网络宽度提高多尺度特征提取的能力。增加宽度能在一定程度上提高效率,但过度增加宽度会出现过拟合等问题。残差神经网络(ResNet)^[14] 通过使用残差模块的跨层连接有效地解决了网络梯度消失和梯度爆炸的问题,使得网络的层数可以达到上千层。虽然之后有更加复杂的网络拥有更高的识别准确率,但是很多下游应用仍将 ResNet 作为自己的基础网络,如物体检测^[15],以及实例分割等^[16],同时在同等复杂度的网络模型中 ResNet 表现得更好。

鉴于上述背景,本文提出一种忆阻器卷积神经网络的表情识别方法。首先基于忆阻器构建了 ResNet 神经网络,并对 ResNet 网络进行剪枝操作,极大地降低了忆阻器神经网络的尺寸以及能耗。然后将 ResNet 模型的所有卷积层以及全连接层的权重映射为忆阻器十字交叉阵列中忆阻器的电导值,用忆阻器十字交叉阵列替代卷积层以及全连接层。实验结果显示忆阻器卷积神经网络模型在 FER2013 数据集上的识别准确率为 63.82%,在 CK+ 数据集上的识别准确率为 93.95%。本文的贡献主要包括:1)首次提出通过忆阻器神经网络来识别面部表情。2)相比其它忆阻器神经网络,本文使用忆阻器搭建的 ResNet 网络复杂度更高,并且广泛应用于各种应用场景中,可以解决实际问题,因此更具有研究价值。3)研究了忆阻器在实际应用中的非理想器件特性对于准确率的影响,为忆阻器神经网络的实际部署提供参考。

1 神经网络改进

1.1 Dropout

在边缘计算设备上应用存储空间大、计算资源消耗大

的深度学习模型是困难的,为了减少计算空间和时间,采用模型压缩方法。在深度学习中,可以通过 Dropout^[17] 来实现模型压缩,具体操作如图 1 所示。

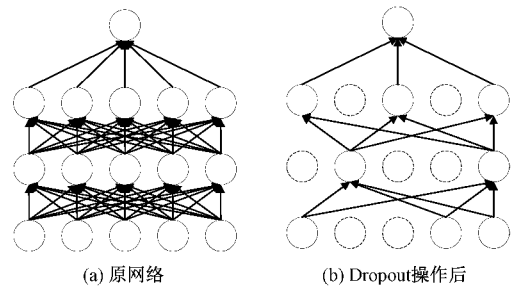


图 1 Dropout 操作

图 1(b)中未连接的节点无需花费忆阻器搭建^[18],这一操作可以实现忆阻器神经网络中的网络剪枝,在提升模型泛化能力的同时减少神经网络中的权重个数,极大地减少搭建神经网络所使用的忆阻器个数,降低模型的物理尺寸以及能耗,符合边缘设备空间以及功耗受限的情况。

1.2 注意力机制

对于输入特征图,卷积注意力模块(CBAM)^[19] 会沿着通道和空间两个维度生成各自的注意力图,具体结构如图 2 所示。

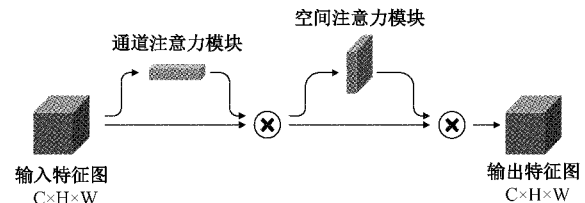


图 2 卷积注意力模块

通道注意力模块如图 3 所示。输入特征图大小为 $C \times H \times W$,分别经过一个全局平均池化和全局最大池化后变为 $C \times 1 \times 1$,之后经过一个由多层感知机和一个隐藏层组成的共享网络中,再对两个向量进行求和,最后经过 sigmoid 激活操作得到输出向量。如图 2 所示,将输出向量和输入特征图做乘法操作,得到通道注意力模块的输出特征图。

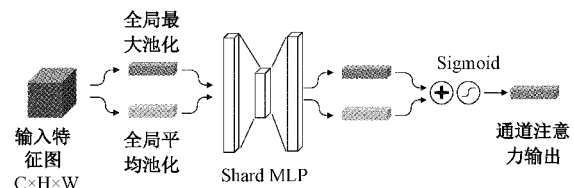


图 3 通道注意力模块

空间注意力模块如图 4 所示。将通道注意力模块输出的特征图作为输入特征图,对输入特征图做一个基于通道的全局平均池化和全局最大池化,通过一个标准的卷积层将它们连接并卷积,再经过 sigmoid 生成空间注意力图。

如图 2 所示,将空间注意力图和空间注意力模块的输入特征图做乘法操作,得到注意力模块的输出特征图。本文通过实验证明,添加注意力模块在 FER2013 上的准确率提升 0.51%,在 CK+ 上的准确率提升 0.75%。

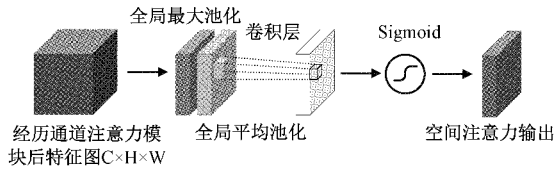


图 4 空间注意力模块

1.3 迁移学习

迁移学习^[20]是指将用来解决任务 A 模型的网络结构和权值用来解决任务 B。实现的原因在于卷积神经网络在浅层学习的特征中存在通用性。迁移学习的优越性在数据不足的情况下尤为突出,当数据集规模较小时,使用迁移学习具有较好的效果^[21]。本文使用在 ImageNet^[22]上训练好的 1 000 分类 ResNet-18 模型,将训练好的模型参数加载到未训练的模型中,再对模型进行训练,从而实现迁移学习。

1.4 ResNet

残差模块是 ResNet 网络的核心^[23],如图 5 所示,在输入和输出之间添加恒等映射,将本来的学习目标 $H(x)$ 转换为残差 $F(x)$,即:

$$F(x) = H(x) - x \tag{1}$$

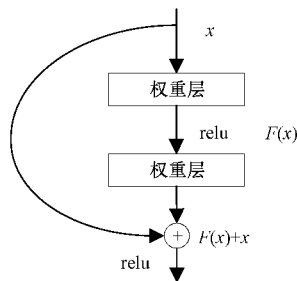


图 5 残差模块

图 6 为本文所使用的残差模块,首先是一个 3×3 的卷积层,之后是 BatchNorm 层,再叠加一个 ReLU 层,之后再叠加一个卷积层和 BatchNorm 层,最后将输入与输出做一个短接操作,得到完整的残差模块。

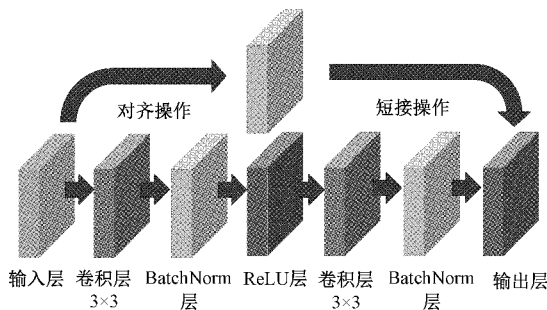


图 6 残差模块网络结构

本文在 ResNet-18 的基础上提出改进,在预处理层与最后一个残差模块之后分别添加了一个注意力模块。为了减少该电路的层数和功耗,将 Dropout 合并到忆阻器神经网络的前几层中。假设输入图片为 $224 \times 224 \times 3$,第一层先经过步长为 2 的 7×7 的卷积,同时进行补零操作,将通道数从 3 上升为 64,再接上一个 BatchNorm 层和 ReLU 层,最后添加一个 Dropout 层完成预处理操作,预处理层的网络结构如图 7 所示。

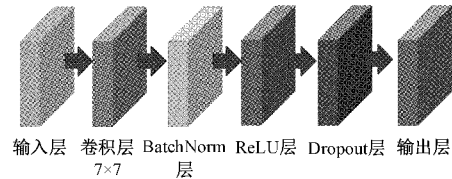


图 7 预处理层网络结构

预处理之后添加一个注意力模块和最大池化层,再连上 4 个卷积步长为 2 的残差模块,每个模块的通道数较之前模块上升一倍,之后再添加一个注意力模块和全局平均池化层,最后通过全连接层得出分类结果,完整的网络结构如图 8 所示。

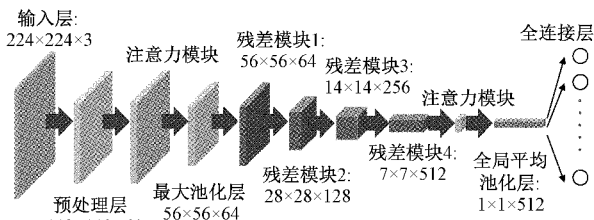


图 8 ResNet-18 网络结构

2 忆阻器卷积神经网络

2.1 VTEAM 忆阻器模型

忆阻器概念是由蔡少棠教授于 1971 年提出的^[24],2008 年惠普实验室第一次实现了一种荷控型忆阻器^[25]。忆阻器阻值由流过忆阻器电荷的多少来改变,因此可以通过控制忆阻器输入电流或输入电压来改变其阻值。压控型忆阻器伏安关系^[26]可描述为:

$$i(t) = G(w, v) \cdot v(t) \tag{2}$$

$$\frac{dw(t)}{dt} = f(w, v) \tag{3}$$

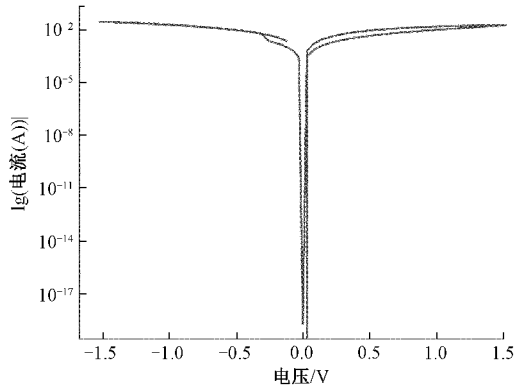
其中, $v(t)$ 是忆阻器两端的电压值, $i(t)$ 是通过忆阻器件的电流, w 是内部的状态变量, $G(w, v)$ 是忆阻器的电导。文献^[27]提出了一种电压阈值自适应忆阻器 (voltage threshold adaptive memristor, VTEAM) 模型,只有当所施加电压的绝对值大于某一阈值时,忆阻器的状态才发生。VTEAM 模型的电压-电流关系未作具体定义,可从任何电压-电流特性中自由选择。VTEAM 实现模型的状态变量导数为:

$$\frac{d\omega(t)}{dt} = \begin{cases} k_{off} \cdot \left(\frac{v(t)}{v_{off}} - 1\right)^{\alpha_{off}} \cdot f_{off}(\omega), & 0 < v_{off} < v \\ 0, & \text{其他} \\ k_{on} \cdot \left(\frac{v(t)}{v_{off}} - 1\right)^{\alpha_{on}} \cdot f_{on}(\omega), & v < v_{on} < 0 \end{cases} \quad (4)$$

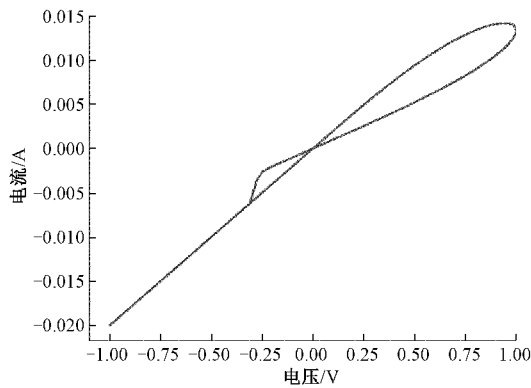
其中, v_{on} 、 v_{off} 表示的是阈值电压, k_{on} 、 k_{off} 、 α_{on} 、 α_{off} 都是常量。参数 k_{off} 是正数, k_{on} 是负数。函数 $f_{off}(\omega)$ 和 $f_{on}(\omega)$ 的作用类似于窗口函数, 目的是保证状态变量 $\omega \in [\omega_{on}, \omega_{off}]$, ω_{off} 和 ω_{on} 分别是内部状态变量 ω 的上下界。当 ω 不在这个范围内时 ω 为 0。忆阻器的电流-电压关系可获得为:

$$i(t) = \left[R_{on} + \frac{R_{off} - R_{on}}{\omega_{off} - \omega_{on}} \cdot (\omega - \omega_{on}) \right]^{-1} \cdot v(t) \quad (5)$$

其中, R_{on} 和 R_{off} 表示状态变量为 ω_{on} 和 ω_{off} 时忆阻器的阻值。对上述 VTEAM 模型进行仿真, 所得到的双极电阻开关特性和磁滞回线如图 9 所示, 输入信号为幅值为 1、频率为 5×10^4 kHz 的正弦信号, $R_{on} = 50 \Omega$, $R_{off} = 1 \text{ k}\Omega$, $v_{on} = -0.2 \text{ V}$, $v_{off} = 0.02 \text{ V}$, $\alpha_{on} = 3$, $\alpha_{off} = 1$, $k_{on} = -10 \text{ m/s}$, $k_{off} = 5 \times 10^{-1} \text{ m/s}$, $\omega_{on} = 0 \text{ m}$, $\omega_{off} = 3 \times 10^{-9} \text{ m}$ 。



(a) 双极电阻开关特性



(b) 磁滞回线

图 9 VTEAM 仿真结果

2.2 1T1R 忆阻器单元

图 10(a) 显示了 1-Transistor 1-Resistor (1T1R) 忆阻器单元^[28], 该单元有 3 个端子连接到十字交叉阵列的源线

(source line, SL)、字线 (word line, WL) 和位线 (bit line, BL)。忆阻器单元的操作取决于存取电流的持续时间以及施加在 WL, BL 和 SL 上的电源电压值。图 10(b) 显示了执行写入操作时的偏置条件。如果执行写入 1 操作, V_{w1} 应用于 WL, BL 接地, 从而将忆阻器的电阻改变为高电阻状态。如果执行写入 0 操作, V_{w0} 应用于 BL, WL 接地, 从而将忆阻器的电阻改变为低电阻状态。从选定的忆阻器单元中读取信息的操作是将一个小于阈值电压的读取电压 V_{read} 连接到 BL 上, 将 WL 接地, 如图 10(c) 所示。因为读取电压小于阈值电压的原因, 所以读取阶段不影响忆阻器阻值。

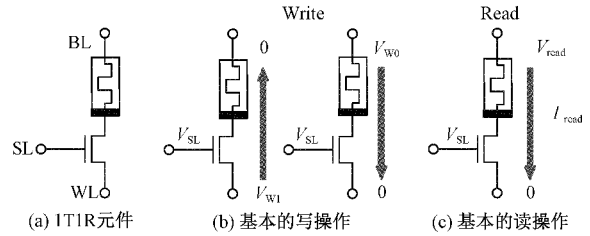


图 10 1T1R 结构

通过 1T1R 结构, 可以控制十字交叉阵列内单个忆阻器的电导值。这允许通过十字交叉阵列中的每个晶体管隔离电流路径, 以便将每个重量编程为目标值。

2.3 忆阻器卷积神经网络

忆阻器特性决定了它的阻值状态是无限的, 因此可以用来表示图像的灰度值。忆阻器卷积神经网络的本质是用忆阻器十字交叉阵列来替代卷积神经网络中的卷积层以及全连接层, 用忆阻器的电导值来表示网络权重。本文通过 MemTorch^[29] 来搭建上述忆阻器卷积神经网络。由于忆阻器不能实现负电导, 因此在忆阻器卷积神经网络中, 权重矩阵可以使用每层两个忆阻器单元来获得正负值。若 g_{pos} 和 g_{neg} 分别表示正权重和负权重^[30], 则单个权重 g_m 为:

$$g_m = g_{pos} - g_{neg} \quad (6)$$

1T1R 结构的忆阻器十字交叉阵列的结构如图 11 所示。

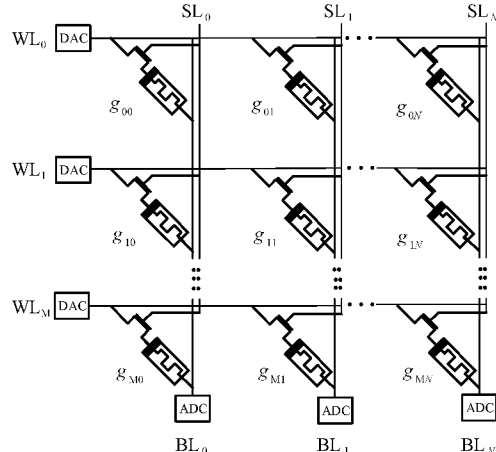


图 11 1T1R 结构的忆阻器十字交叉阵列

用电导 g 来代表权重矩阵 W , 其中 $[WL_0, WL_1, WL_2, \dots, WL_M]$ 表示输入的信号, 通过数模转换器将特征图片的像素信息通过等比例缩放转化为输入电压, $[g_{00}, g_{01}, \dots, g_{0N}, g_{10}, \dots, g_{1N}, \dots, g_{M0}, \dots, g_{MN}]$ 表示在十字交叉阵列内的忆阻器电导, $[BL_0, BL_1, BL_2, \dots, BL_N]$ 表示的是输出电流, 忆阻器十字交叉阵列的输入输出关系如式(7)所示。

$$[BL_0 \ \dots \ BL_N] = [WL_0 \ \dots \ WL_M] \begin{bmatrix} g_{00} & \dots & g_{0N} \\ \dots & \dots & \dots \\ g_{M0} & \dots & g_{MN} \end{bmatrix} \quad (7)$$

通过交叉阵列的 SL 控制忆阻器的开断, 与每个十字交叉阵列的单个模数转换器结合使用, 可以通过时分复用的方式按顺序读取每列的电流, 实现对忆阻器十字交叉阵列中单个忆阻器的读写操作。

忆阻器十字交叉阵列的卷积过程如图 12 所示, 假设输入为一个单通道的 3×3 的图片, 有两个 2×2 的卷积核, 分别对输入图像进行卷积操作。以第一次卷积操作为例, 首先要将输入图片进行卷积操作的部分 $[D_0, D_1, D_3, D_4]$ 进行展开操作, 同时将两个卷积核分别展开为 $[F_0, F_1, F_2, F_3]$ 和 $[H_0, H_1, H_2, H_3]$, 之后进行卷积操作分别得到输出 $[O_{00}, O_{10}, O_{20}, O_{30}]$ 和 $[O_{01}, O_{11}, O_{21}, O_{31}]$ 。本文通过忆阻器十字交叉阵列完成上述卷积操作, $[D_0, D_1, D_3, D_4]$ 对应的是输入电压值 $[WL_0, WL_1, WL_2, WL_3]$, 通过对图片的像素值进行等比例缩放, 得到忆阻器十字交叉阵列的输入电压。 $[F_0, F_1, F_2, F_3]$ 对应忆阻器十字交叉阵列中的电导值 $[g_{00}, g_{01}, \dots, g_{0N}, g_{10}, \dots, g_{1N}, \dots, g_{M0}, \dots, g_{MN}]$, 输出 $[O_{00}, O_{10}, O_{20}, O_{30}]$ 对应忆阻器十字交叉阵列中的输出电流 $[BL_0, BL_1, BL_2, BL_3]$ 。

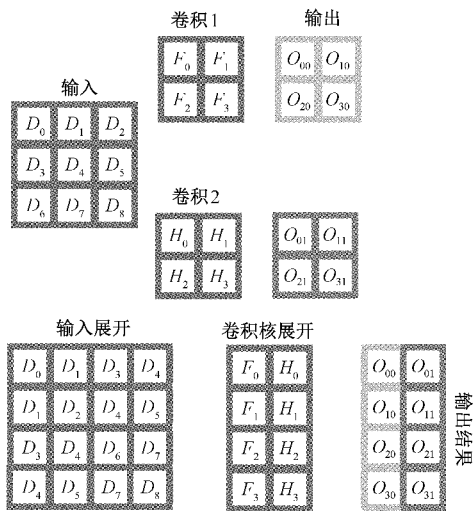


图 12 忆阻器十字交叉阵列卷积过程

2.4 忆阻器的非理想器件特性

虽然理想的忆阻器十字交叉阵列被认为可以显著加速卷积神经网络的学习速度, 大大降低其功耗。但在实际的

架构中, 当交叉阵列尺寸被放大时, 实验制造的忆阻器器件缺陷会导致其性能显著下降。这些缺陷包括非线性特性、随机电导更新、器件质量随时间和空间变化、器件成品率以及有限数量的电导状态^[31]。为了尽量减少这些缺陷的影响, 可以使用特定的外围电路和系统缓解技术, 不过这些技术增加了系统计算的时间复杂性。了解忆阻器非理想器件特性的影响可以更好地优化整体性能, 因此, 在将忆阻器十字交叉阵列用于任何应用之前, 必须考虑到这些非理想特性的影响, 其中对准确性的影响至关重要。由于生产工艺的限制, 现实的忆阻器器件是非理想的, 具有有限数量稳定离散的可切换电导状态, 其中电导状态由 R_{on} 和 R_{off} 决定^[32]。忆阻器极易发生故障, 要么无法进行电阻状态的更新, 要么被卡在高电阻或低电阻状态^[33]。本文主要研究有限数量的电导状态以及设备故障这两个非理想器件特性对于准确率的影响, 目的是模拟忆阻器在实际应用中的真实情况, 为以后忆阻器的应用发展提供参考。

3 表情识别实验

3.1 数据集

选择的数据集是 FER2013、CK+数据集, 这两个数据集的使用较为广泛, 具有一定的参考价值^[34]。

FER2013 是由 International Conference On Machine Learning (ICML)2013 挑战产生, 包含了 35 887 张图片, 其中七种表情分别为 Angry、Disgust、Fear、Happy、Sadness、Surprise 和 Neutral。

CK+数据集是对 123 个实验者拍摄了 593 视频序列, 这些序列展示了测试对象的从无表情到有表情的转变, 选择每个视频序列中具有明显表情特征的最后三帧作为实验对象, 同时将剩下的归到无表情 Neutral 类, 其中八种表情标签分别为 Angry、Contempt、Disgust、Fear、Happy、Sadness、Surprise 和 Neutral。

FER2013 与 CK+相比, 识别难度较高, 如文献[10]所示, 在使用同样的方法对 FER2013 与 CK+进行识别时, FER2013 的识别准确度与 CK+相比低了 29.83%。主要原因是 FER2013 数据集本身存在问题, 其中包含卡通人物, 水印遮挡等难以分类的样本, 同时 FER2013 的训练集和测试集图片样本差距较大, 进一步提高了测试集的识别难度, 因此本文在 FER2013 与 CK+这两个不同数据集上的准确率表现有不同。

3.2 实验流程

实验环境如下, 使用的操作系统是 ubuntu20.04, Python 版本 3.7.3, PyTorch 版本 1.1.0, 显卡型号 RTX3070。

首先用 FER2013 数据集和 CK+数据集分别对 ResNet-18 网络进行训练, 其中 FER2013 的训练轮数为 200, CK+的训练集的训练轮数为 1 000, 初始学习率为 $\eta = 1 \times 10^{-3}$ 。对 FER2013, 每经过 30 轮学习率减少为之前的

一半,而对 CK+, 每经过 100 轮学习率减少为之前的一半。通过 CrossEntropyLoss 函数确定网络的损失函数,通过 Adam 算法对网络参数进行优化,得出训练好的 ResNet-18 模型,最后通过测试集对模型的准确率进行测试。

通过 MemTorch 将 ResNet 模型的所有卷积层以及全连接层的权重映射为忆阻器十字交叉阵列中忆阻器的电导值,用忆阻器十字交叉阵列替代卷积层以及全连接层,构建出忆阻器卷积神经网络。具体操作就是将原模型的卷积层以及全连接层通过 MemTorch 提供的应用程序接口进行重新编码,得到新的忆阻器神经网络模型。同时将忆阻器神经网络的输入缩放为 $-0.3 \sim +0.3$, 对应到模拟电路上就是将电压控制在 $-0.3 \sim +0.3$ V。所使用的 VTEAM 忆阻器模型中取 $R_{on} = 50 \Omega, R_{off} = 1 \text{ k}\Omega$, 使用大小为 128×128 的十字交叉阵列。之后对忆阻器卷积神经网络的模型进行准确率测试,并且与原网络模型准确率进行对比。最后对忆阻器的非理想特性进行建模,测试其对忆阻器神经网络准确率的影响。

3.3 CK+实验结果

训练之前需对数据集做一些预处理操作。首先是对数据集做一个统一的 resize 操作,将所有的图片都 resize 成 224×224 。之后通过随机旋转,随机切割,从而扩大数据集的规模实现数据增强。最后对数据进行归一化处理,将所需图片像素数值范围从 0 的一侧 $0 \sim 1$ 转化为在 0 的两侧分布 $-0.5 \sim 0.5$, 使得训练过程更加稳定,同时加速模型在训练时的收敛速度。归一化所使用的 RGB 通道均值分别为 $(0.458, 0.456, 0.406)$, 方差分别为 $(0.229, 0.224, 0.225)$, 这些数值是从 ImageNet 数据集中统计出来的, ImageNet 数据集具有几百万张图片,具有一定代表性,图 13 是预处理之后图片的样子。



图 13 CK+部分提取

最后的识别准确率达到到了 95.19%,图 14 所示为用 ResNet-18 训练 1 000 轮得到的准确率和损失函数。

图 14(a) 中横坐标表示训练的批次,纵坐标表示 ResNet 模型的识别准确率,即识别图片正确表情的样本数占总样本数的比例。图 14(b) 表示训练过程中交叉熵损失函数值的变化,图中横坐标表示训练的批次,纵坐标表示该批次损失函数的具体值。

实验显示基于忆阻器卷积神经网络表情识别准确率达

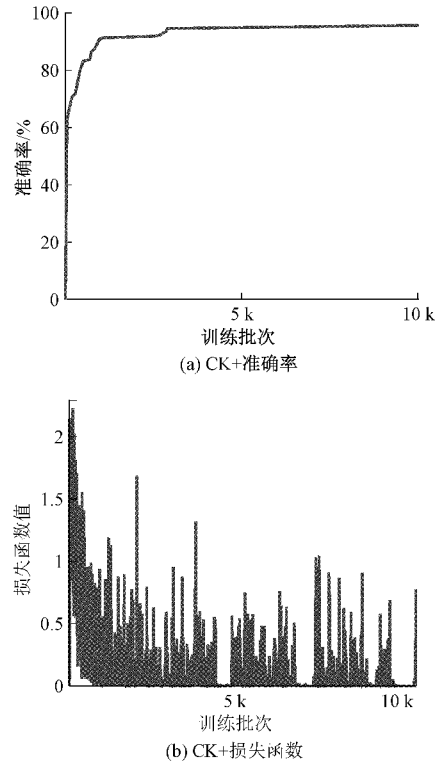


图 14 CK+实验结果

到了 94.43%, 相比原卷积神经网络的 95.19% 准确率损失为 0.76%, 充分证明忆阻器搭建的复杂神经网络具有完成表情识别这一复杂分类任务的能力。

对于有限个电导状态忆阻器的实验结果,将忆阻器的电导个数从无限个设置为有限个,取 $10 \sim 20$ 十一组有限个电导状态。由有限个电导状态忆阻器搭建的神经网络在测试集上的准确率表现如图 15 所示,其中理想状态忆阻器搭建的神经网络在测试集上的准确率为 94.99%,有限个电导状态下准确率最低值为 15 个电导状态下的 93.95%,证明与理想状态忆阻器相比,有限个电导状态的忆阻器会对准确率产生影响,同时具有不同个数的电导状态的忆阻器对准确率的影响程度不同。

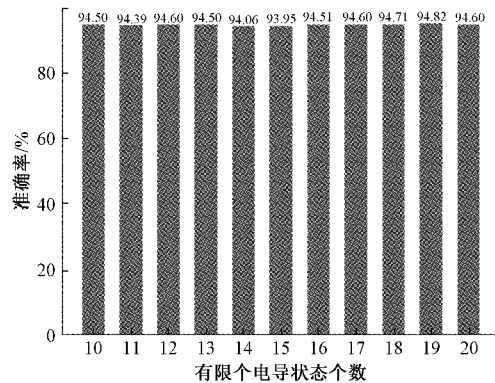


图 15 CK+中有限个电导状态下忆阻器神经网络模型准确率对比

对于忆阻器设备故障对准确率的影响,主要是对卡在高电阻或低电阻状态这一故障进行研究,通过 3 组数据进行对比。第 1 组是高电阻和低电阻状态都发生故障,第 2 组是低电阻状态发生故障,第 3 组是高电阻状态发生故障,3 组的实验数据如表 1 所示。将每一组的数据做一个对比,发现当高低电阻同时发生故障时,准确率与无故障时的准确率相差最大。故障率是指忆阻器发生故障的个数占总体个数的比例,当故障率达到 0.225 时的准确率降低了 14.46%。当低电阻发生故障时,准确率也会有明显的下降,当故障率为 0.25 时的准确率下降率降低了 9.08%。高电阻故障时对于准确率的影响较低,当故障率为 0.225 时准确率下降率 7.63%。

表 1 忆阻器故障对准确率的影响 %

故障率	高低电阻故障准确率	低电阻故障准确率	高电阻故障准确率
0.000	94.58	94.04	94.58
0.025	94.58	93.58	94.04
0.050	93.67	93.04	94.21
0.075	93.50	93.04	91.50
0.100	92.95	91.41	92.67
0.125	94.13	91.13	89.67
0.150	91.04	92.13	88.04
0.175	82.26	91.67	89.02
0.200	85.13	90.04	88.56
0.225	80.12	85.28	86.95
0.250	82.16	84.96	90.21

3.4 FER2013 实验结果

FER2013 的数据集在训练之前进行与 CK+ 同样的预处理操作,图 16 是预处理之后图片的样子。

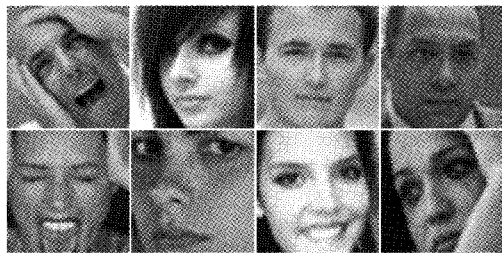
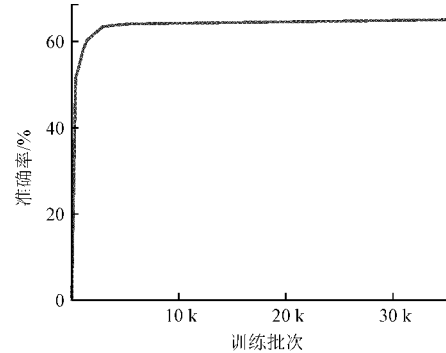


图 16 FER2013 预处理之后的部分提取

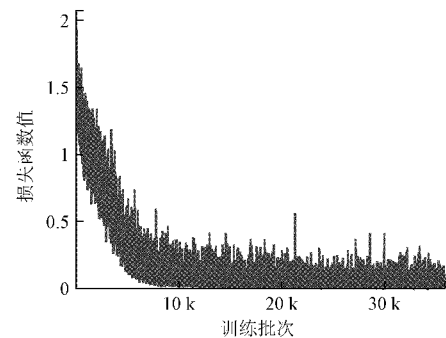
FER2013 的最后的识别准确率达到 64.13%,用 ResNet-18 训练 200 轮得到的准确率和损失函数如图 17 所示。

实验显示忆阻器卷积神经网络识别准确率达到 63.82%,相比原卷积神经网络的 64.13% 准确率损失为 0.31%。

对于有限个电导状态忆阻器的实验结果,将忆阻器的



(a) FER2013 准确率



(b) FER2013 损失函数

图 17 FER2013 实验结果

电导个数从无限个设置为有限个,取 10~20 十一组不同的电导状态,由有限个电导状态忆阻器搭建的神经网络在测试集上的准确率表现如图 18 所示,其中理想状态忆阻器搭建的神经网络在测试集上的准确率为 63.89%,有限个电导状态下准确率最低值为 11 个电导状态下的 62.76%。

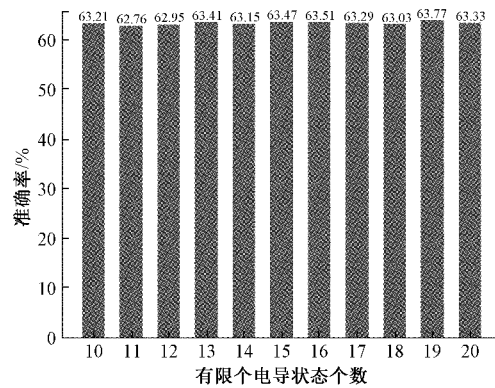


图 18 中有限个电导状态下忆阻器神经网络模型准确率对比

对于忆阻器设备故障对准确率的影响,3 组的实验数据如表 2 所示。将每一组的数据做一个对比,发现当高低电阻同时发生故障时,准确率与无故障时的准确率相差最大,当故障率达到 0.25 时的准确率降低了 32.49%。当低电阻发生故障时,准确率也会有明显的下降,当故障率为 0.225 时的准确率下降率降低了 17.90%。高电阻故障时对于 FER2013 的准确率的影响较低。

表 2 忆阻器故障对准确率的影响 %

故障率	高低电阻故障准确率	低电阻故障准确率	高电阻故障准确率
0.000	63.51	63.19	62.83
0.025	59.90	60.89	63.36
0.050	61.06	62.19	63.69
0.075	56.53	58.21	63.06
0.100	56.89	57.49	62.77
0.125	40.33	51.89	63.56
0.150	37.97	56.56	63.51
0.175	53.66	55.86	63.03
0.200	43.31	57.78	62.98
0.225	49.35	45.29	62.90
0.250	31.02	52.47	62.94

4 结 论

为了解决深度学习模型在边缘设备上的部署问题,本文提出了用忆阻器神经网络来解决的思路。对于人脸情感识别问题,通过对 ResNet-18 网络进行改进,将训练完模型的卷积层以及线性层的权值映射成忆阻器电导值,构成忆阻器卷积神经网络,实验结果证明忆阻器卷积神经网络表情识别仍具有很高的准确性。此外,本文还验证了有限个电导状态以及忆阻器的故障率两种非理想特性对准确率的影响,为忆阻器的实际部署提供参考。

参考文献

- [1] KRESTINSKAYA O, JAMES A P, CHUA L O, et al. Neuromemristive circuits for edge computing: A review[J]. IEEE Transactions on Neural Networks and Learning Systems 2020,31(1):4-23.
- [2] KHELIFI H, LUO S, NOUR B, et al. Bringing deep learning at the edge of information-centric internet of things [J]. IEEE Communications Letters, 2019, 23(1): 52-55.
- [3] YOUSEFPOUR A, FUNG C, NGUYEN T, et al. All one needs to know about fog computing and related edge computing paradigms: A complete survey[J]. Journal of Systems Architecture, 2019, 98: 289-330.
- [4] ROY K, JAISWAL A, PANDA P. Towards spike-based machine intelligence with neuromorphic computing[J]. Nature, 2019, 575(7784): 607-617.
- [5] WANG Y, CAO Y, GUO Z, et al. Event-based sliding-mode synchronization of delayed memristive neural networks via continuous/periodic sampling algorithm[J]. Applied Mathematics and Computation, 2020, DOI:10.1016/j.amc.2020.125379.
- [6] KATAEVA I, MERRIKH-BAYAT F, ZAMANIDOOST E, et al. Efficient training algorithms for neural networks based on memristive crossbar circuits [C]. 2015 International Joint Conference on Neural Networks (IJCNN), IEEE, 2015: 1-8.
- [7] YAO P, WU H, GAO B, et al. Face classification using electronic synapses[J]. Nature Communications, 2017, 8(1): 1-8.
- [8] YAO P, WU H, GAO B, et al. Fully hardware-implemented memristor convolutional neural network[J]. Nature, 2020, 577 (7792): 641-646, DOI: 10.1038/s41586-020-1942-4.
- [9] 张波,兰艳亭,鲜浩,等.基于通道注意力机制的人脸表情识别机器人交互研究[J].电子测量技术,2021,44(11):169-174.
- [10] 程焕新,成凯,蒋泽芹.基于注意力机制的 CNN 人脸表情识别[J].电子测量技术,2021,44(10):128-132.
- [11] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. Imagenet classification with deep convolutional neural networks [J]. Advances in Neural Information Processing Systems, 2012, 25: 1097-1105.
- [12] SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition [J]. ArXiv Preprint,2014, ArXiv:1409.1556.
- [13] SZEGEDY C, LIU W, JIA Y, et al. Going deeper with convolutions[C]. IEEE Conference on Computer Vision and Pattern Recognition, 2015: 1-9.
- [14] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition[C]. IEEE Conference on Computer Vision and Pattern Recognition, 2016: 770-778.
- [15] REN S, HE K, GIRSHICK R, et al. Faster r-cnn: Towards real-time object detection with region proposal networks[J]. Advances in Neural Information Processing Systems, 2015, 28: 91-99.
- [16] HE K, GKIOXARI G, DOLLÁR P, et al. Mask R-CNN [C]. IEEE International Conference on Computer Vision, 2017: 2961-2969.
- [17] SRIVASTAVA N, HINTON G, KRIZHEVSKY A, et al. Dropout: A simple way to prevent neural networks from overfitting[J]. The Journal of Machine learning Research,2014, 15(1): 1929-1958.
- [18] 王立刚,张志佳,李晋,等.基于卷积神经网络的 LED 灯类字体数字识别[J].电子测量与仪器学报,2020,34(11):148-154.
- [19] WOO S, PARK J, LEE J Y, et al. Cbam: Convolutional block attention module[C]. European Conference on Computer Vision(ECCV), 2018: 3-19.
- [20] LONG M, WANG J, CAO Y, et al. Deep learning of transferable representation for scalable domain

- adaptation[J]. *IEEE Transactions on Knowledge and Data Engineering*, 2016, 28(8): 2027-2040.
- [21] 莫宏伟,傅智杰. 基于迁移学习的无监督跨域人脸表情识别[J]. *智能系统学报*, 2021, 16(3): 397-406.
- [22] DENG J, WEI D, RICHARD S, et al. Imagenet: A large-scale hierarchical image database [C]. *IEEE Conference on Computer Vision and Pattern Recognition*, 2009: 248-255.
- [23] 韩冉,曾广森,王荣杰. 基于残差块网络的图像去雨算法[J]. *仪器仪表学报*, 2021, 41(8): 175-182.
- [24] CHUA L. Memristor-the missing circuit element[J]. *IEEE Transactions on Circuit Theory*, 1971, 18(5): 507-519.
- [25] STRUKOV D B, SNIDER G S, STEWART D R, et al. The missing memristor found [J]. *Nature*, 2008, 453 (7191): 80-83, DOI: 10. 1109/TCSII. 2015. 2433536.
- [26] CHUA L O, KANG S M. Memristive devices and systems[J]. *Proceedings of the IEEE*, 1976, 64(2): 209-223.
- [27] KVATINSKY S, RAMADAN M, FRIEDMAN E G, et al. VTEAM: A general model for voltage-controlled memristors[J]. *IEEE Transactions on Circuits and Systems II: Express Briefs*, 2015, 62(8): 786-790.
- [28] CHEN Y X, LI J F. Fault modeling and testing of 1T1R memristor memories[C]. 2015 IEEE 33rd VLSI Test Symposium(VTS), IEEE, 2015: 1-6.
- [29] LAMMIE C, AZGHADI M R. MemTorch: A simulation framework for deep memristive Cross-Bar architectures[C]. 2020 IEEE International Symposium on Circuits and Systems(ISCAS), IEEE, 2020: 1-5.
- [30] ALIBART F, ZAMANIDOOST E, STRUKOV D B. Pattern classification by memristive crossbar circuits using ex situ and in situ training [J]. *Nature Communications*, 2013, 4(1): 1-7.
- [31] BURR G W, NARAYANAN P, SHELBY R M, et al. Large-scale neural networks implemented with non-volatile memory as the synaptic weight element: Comparative performance analysis (accuracy, speed, and power) [C]. 2015 IEEE International Electron Devices Meeting(IEDM), IEEE, 2015.
- [32] YI W, SABEL'EV SE, MEDEIROS R G, et al. Quantized conductance coincides with state instability and excess noise in tantalum oxide memristors[J]. *Nature Communications*, 2016, 7(1): 1-6.
- [33] MEHONIC A, JOKSAS D, NG W H, et al. Simulation of inference accuracy using realistic RRAM devices[J]. *Frontiers in Neuroscience*, 2019, 13, DOI: 10. 3389/fnins. 2019. 00593.
- [34] 李珊,邓伟洪. 深度人脸表情识别研究进展[J]. *中国图象图形学报*, 2020, 25(11): 2306-2320.

作者简介

赵益波,副教授,博士,主要研究方向为非线性电路分析与设计、自适应信号处理与忆阻器神经网络。

E-mail: yibozaodn@163.com

蒋文,硕士,主要研究方向为忆阻器神经网络应用与深度学习。

E-mail: 1600474822@qq.com

孟若禹,硕士,主要研究方向为忆阻器神经网络应用与深度学习。

E-mail: 595948447@qq.com

李业宁,硕士,主要研究方向为忆阻器神经网络应用与深度学习。

E-mail: 1547811550@qq.com