

DOI: 10.19650/j.cnki.cjsi.J2210720

基于振动加速度信息的触觉纹理分类方法研究*

陈大鹏^{1,2,3}, 朱栋梁^{1,2}, 刘佳^{1,2,3}, 宋爱国^{2,4}, 陈庚^{1,2}

(1. 南京信息工程大学自动化学院 南京 210044; 2. 江苏省智能气象探测机器人工程研究中心 南京 210044;
3. 江苏省大气环境与装备技术协同创新中心 南京 210044; 4. 东南大学仪器科学与工程学院 南京 210096)

摘要:当手持刚性工具在材料表面滑动时,用户可以通过工具的振动来感受材料表面的纹理特征。这些振动加速度数据包含了丰富的纹理类别信息,为纹理的分类提供了基础。利用触觉进行纹理分类对于力触觉人机交互、机器人精细化操作等应用具有重要的意义。目前,手工设计与纹理相关的特征以及借助卷积神经网络进行简单的特征提取等方法已经被应用于触觉纹理分类。然而,这些方法未能关注时间尺度的选择和触觉序列数据间的时间依赖性,还存在触觉数据特征提取不充分和分类精度不佳等问题。为了解决上述问题,本文提出一种由多尺度卷积网络和双向长短时记忆网络相结合的融合模型,以便同时捕获触觉信号多尺度的几何局部空间特征和时间依赖特征。所提出的模型从公开的触觉数据集中学习材料表面纹理的触觉特征,并在公开的纹理振动加速度数据库上进行训练。实验结果表明,本文提出的模型可以稳健且高效地实现最高 92.1% 的纹理分类精度。

关键词: 触觉纹理分类;深度学习;振动加速度数据

中图分类号: TH7 TP391 **文献标识码:** A **国家标准学科分类代码:** 520.60

Research on the tactile texture classification method based on vibration acceleration information

Chen Dapeng^{1,2,3}, Zhu Dongliang^{1,2}, Liu Jia^{1,2,3}, Song Aiguo^{2,4}, Chen Geng^{1,2}

(1. School of Automation, Nanjing University of Information Science & Technology, Nanjing 210044, China; 2. Jiangsu Province Engineering Research Center of Intelligent Meteorological Exploration Robot (C-IMER), Nanjing 210044, China; 3. Jiangsu Collaborative Innovation Center of Atmospheric Environment and Equipment Technology (CICAET), Nanjing 210044, China; 4. School of Instrument Science and Engineering, Southeast University, Nanjing 210096, China)

Abstract: When a user holds a rigid tool to slide on the material surface, the texture features of the material surface through vibration of the tool is felt. These vibration acceleration data contain rich texture category information, which provides a basis for texture classification. Texture classification based on tactile sense is of great significance for applications such as haptic human-computer interaction and fine manipulation of robots. At present, the methods of manually designing features related to texture and simple feature extraction using convolutional neural network have been applied to tactile texture classification. However, these methods fail to pay attention to the selection of time scale and the time dependence between tactile serial data, and there are still problems such as insufficient feature extraction of tactile data and poor classification accuracy. To solve the above problems, this article proposes a fusion model which combines multi-scale convolutional network and bidirectional long short memory network to capture multi-scale geometric local spatial features and time dependent features of tactile signals at the same time. The proposed model learns the tactile features of material surface texture from an open tactile data set, and trains them on the open texture vibration acceleration database. The experimental results show that the proposed model achieves the highest texture classification accuracy of 92.1% robustly and efficiently.

Keywords: tactile texture classification; deep learning; vibration acceleration data

收稿日期: 2022-11-16 Received Date: 2022-11-16

* 基金项目: 国家自然科学基金(62003169, 61773219)、江苏省自然科学基金青年基金(BK20200823)、江苏省高等学校自然科学研究项目(20KJB520029)、江苏省研究生科研与实践创新计划项目(SJCX22_0348)资助

0 引 言

纹理是物体表面普遍存在且具有视-触觉双重属性的一种基本特征^[1]。通过对纹理进行提取和分析,纹理分类已经在物体识别、人脸识别、图像和视频检索、遥感/医学图像分析等计算机视觉和模式识别领域得到广泛的应用。目前的纹理分类研究主要是根据纹理中包含的视觉线索并利用图像处理算法完成的^[2]。除了基于视觉图像的纹理分类方法,触觉也能用于对纹理进行分类。触觉纹理分类是用户或机器根据与材料表面交互时的力/触觉反馈信息对纹理进行辨识和区分的技术。对于一些特定的应用场景,如盲人识别物体、力触觉人机交互、机器人触觉、虚拟手术训练等,触觉纹理往往发挥着至关重要且难以被视觉纹理替代的作用^[3-5]。同时,在虚拟/增强现实等交互场景中,给虚拟物体添加上触觉纹理感受还可以显著增强人机交互的真实感和临场感^[6]。

纹理包含了物体的粗糙度、硬度、摩擦等多种触觉感知维度,能够反映物体表面状态的多重信息^[7]。当用户使用刚性工具探索纹理表面时,工具上产生的振动加速度数据与纹理的特征信息高度相关^[8]。这些振动加速度数据潜在地可以被用于对纹理进行分类。例如,Jamali 等^[9]在工业机器人的手臂上安装模仿人类手指的仿生传感器用于探索表面纹理,为纹理分类提供振动加速度数据,然后利用谱包络的 5 个主峰形成特征向量。该方法在匹配/分类时的特征定义很大程度上取决于纹理探索过程的压力、速度等参数。然而,由于振动加速度信号会随作用力和扫描速度的变化而变化^[10],机器人需要遵循定义良好的轨迹才能对表面进行探索。Strese 等^[8,11]利用音频处理中广泛使用的成熟声学特征对振动加速度数据进行处理,以达到纹理分类的目的。Fishel 等^[12]提出的数据驱动方法与专门设计的数据预处理算法(如阈值处理和峰值检测)一起工作,实现了所部署的感知特征与特定分类任务的紧密耦合。Nie 等^[13]研究了非结构化环境中触觉纹理数据时间窗口的选择对纹理分类的影响,提出了一种确定触觉数据时间窗口范围的方法,并利用基本的机器学习分类来验证特征设计和触觉数据时间窗口上下限选择的有效性。

虽然上述借助机器学习对纹理进行分类的方法已经取得了不错的效果,但是仍然存在一些弊端。比如,先前的方法在对触觉信号进行预处理时,通常需要从原始的触觉数据中找寻规律,并手动设计与触觉相关的纹理特征。这种通过一系列工程化的方式从原始数据中筛选出较好数据特征的方法需要大量的专家经验,无法很好地适用于对未知纹理的分类,且整个过程繁琐复杂。

近年来,深度神经网络因其强大的自主学习能力而被应用于解决这类问题。与以往手动提取特征的方法相比,基于深度学习的方法可以简化预处理的过程并且训练出的模型具备较好的泛化能力。一些研究已经利用深度学习从振动加速度信号这种时间序列数据中直接提取纹理的触觉特征。例如,Ji 等^[14]提出了一种由稀疏自动编码器预训练的卷积神经网络(convolutional neural network, CNN)对原始加速度数据进行纹理分类的方法,避免了预处理阶段繁琐的手工设计特征并且获得了良好的分类效果。Orii 等^[15]使用长短时记忆(long short-term memory, LSTM)网络训练采集到的多种触觉信号,用以识别纹理。Bednarek 等^[16]和惠文珊等^[17]分别基于 CNN 模型与 LSTM 模型串联结合的方式对纹理分类展开研究,他们在自己采集的纹理数据集上得到了较好的分类结果。Joollee 等^[18]提出了一种 SpectroNet、一维残差 CNN 和 Bi-GRU 相结合的多模型融合网络从触觉数据中提取材料的纹理特征。基于无监督学习理论,Metzger 等^[19]训练了深度卷积自编码器将振动信号压缩成 16 个特征的潜在表示,并对其进行重构。上述的研究表明,在触觉纹理分类方面,基于深度学习的方法有着独特的优势,既能够避免传统方法中复杂的特征提取过程,又能获得较好的分类性能。

在上述研究的基础上,本文基于人使用刚性工具在真实材料表面自由探索期间记录的振动加速度信号^[11],进一步研究了材料表面的触觉纹理分类问题。从真实纹理样本中采集到的振动加速度信号是一种典型的时间序列数据,包含着复杂的表面纹理信息,而且不同长度的子序列包含着不同时间尺度的纹理特征信息。因此,本文重点关注触觉序列数据的多尺度特征以及时序相关性特征,利用多尺度卷积网络(multi-scale convolutional neural network, MCNN)和双向长短时记忆(bidirectional long short-term memory, Bi-LSTM)网络并行的深度架构来对触觉纹理信息进行特征提取,以提高对不同纹理的分类性能。在 MCNN 中,本文为 CNN 设置了 3 种大小不同的卷积核,通过产生多个感受域来提取振动加速度数据在不同尺度上的特征。Bi-LSTM 用于获取输入数据时间域隐含的序列依赖关系。接着,本文利用 LMT 触觉数据库^[11]中 9 类纹理的振动加速度数据对网络模型的有效性进行了综合评估。实验结果表明,相比于现有的纹理分类模型,本文提出的基于 MCNN 和 Bi-LSTM 的融合网络取得了更好的纹理分类效果。

1 触觉纹理分类模型的设计

在本文中,模型的输入是一维的振动加速度数据片段,这些数据片段是由 DFT321^[20]算法对原始的三维加

速度数据进行转换而获得的。因此,卷积核和特征图都是一维的。由于一维的振动加速度数据具有高度的时间依赖性,本文采用一维 MCNN 和 Bi-LSTM 网络结合的网络架构来充分提取振动加速度信号的特征。所提出的模型的总体框架如图 1 所示,该图显示了两个模块:MCNN 模块和 Bi-LSTM 网络模块。在该框架中,原始的振动加速度数据采用滑动窗口法将数据切分,得到的数据片段分别送入 MCNN 模块和 Bi-LSTM 模块。MCNN 模块提取数据的多尺度特征,Bi-LSTM 网络模块捕获数据的时序相关性特征,将两种不同的特征融合得到新的特征表示。采用全局平均池化替换全连接层,减少网络的参数量,防止网络的过拟合。融合的特征经过分类层(Softmax)进行最终的分类预测,得到纹理类别信息。

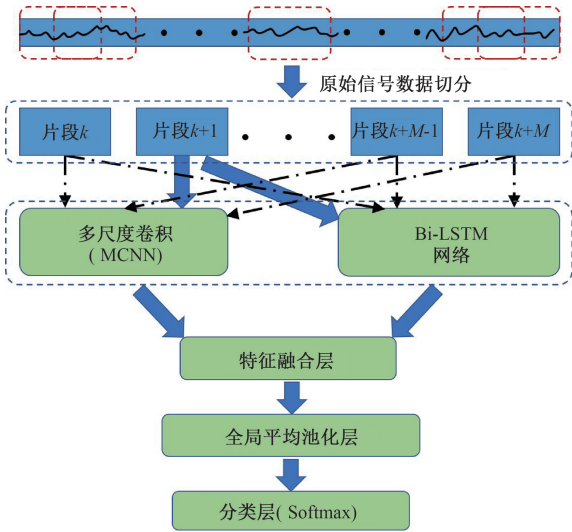


图 1 模型的整体架构

Fig. 1 Overall architecture of the model

1.1 多尺度卷积网络(MCNN)设计

使用刚性工具在真实的纹理表面上以不同的压力和速度滑动时,工具会产生不同频率的振动加速度信号。对于复杂的振动加速度信号,如果使用单尺度的卷积核提取信号的特征,很可能出现特征提取不充分以及数据被错误分类等问题^[21]。因此,单一尺度的卷积核难以应对受多元因素影响的纹理分类问题。

由于振动加速度信号是一种典型的时间序列数据,本文利用能在不同尺度上提取时序数据特征的 MCNN 来取代常规卷积网络中的单一尺度卷积层。通过设置不同大小的卷积核分别对振动加速度数据进行卷积,MCNN 能获得不同视野的感受域,充分挖掘数据的特征,进一步克服局部特征信息提取不充分的不足。MCNN 的结构如图 2 所示,其能够融合丰富的局部信息,并更加充分地利用序列内部的隐含信息。

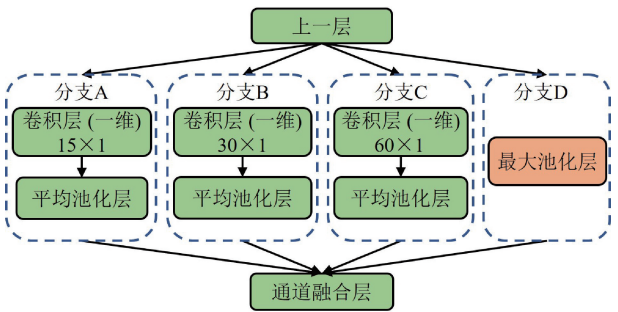


图 2 多尺度卷积网络

Fig. 2 Multi-scale convolution network

图 2 中的网络是由 4 个分支组成,其中分支 A、分支 B 和分支 C 所使用的卷积核大小各不相同,目的是从上一层输出中提取到多尺度的特征。为了对小扰动保持稳定,模型中引入了另一个并行分支 D 用作最大池化。接着,模型对每个分支经卷积/池化操作得到的特征向量进行融合,并将拼接融合的特征向量输入下一层。MCNN 中各个分支所需的卷积核个数如表 1 所示。

表 1 MCNN 中各分支卷积核个数			
Table 1 The number of convolution kernels in each branch			
分支	16×1 卷积核个数	30×1 卷积核个数	60×1 卷积核个数
A	64	—	—
B	—	32	—
C	—	—	16
D	—	—	—

一维多尺度卷积层的计算公式如下:

$$\mathbf{C}^{l+1} = \begin{cases} f(\mathbf{C}^l \mathbf{W}_{16 \times 1}^A + \mathbf{b}_{16 \times 1}^A) \\ f(\mathbf{C}^l \mathbf{W}_{30 \times 1}^B + \mathbf{b}_{30 \times 1}^B) \\ f(\mathbf{C}^l \mathbf{W}_{60 \times 1}^C + \mathbf{b}_{60 \times 1}^C) \end{cases} \quad (1)$$

式中: $\mathbf{C}^l$ 表示上一层的输出激活值; $\mathbf{C}^{l+1}$ 表示下一层的激活值; $\mathbf{W}_{16 \times 1}^A$ 为权重矩阵; $\mathbf{b}_{16 \times 1}^A$ 为偏置系数; f 为激活函数,负责将非线性特性引入网络中。 $\mathbf{W}_{16 \times 1}^A$ 和 $\mathbf{b}_{16 \times 1}^A$ 的上标 A 表示其所归属的分支,下标 16×1 表示卷积核或者偏置矩阵的大小。

1.2 Bi-LSTM 网络设计

Bi-LSTM 网络用于捕获一维触觉纹理数据的时间依赖性,提取触觉信息的时间特征。Bi-LSTM 网络由 LSTM 网络发展而来。LSTM 能从时间序列数据的前向上下文学习中学习特征,但不能学习相同序列数据的后向上下文特征。Bi-LSTM 网络通过使用两个独立的隐藏层在前向和后向两个方向上学习顺序数据,弥补了 LSTM 网络的不足。在数据样本足够多的条件下,Bi-LSTM 在分类性能

上的表现比 LSTM 更好一些^[22]。因此,本文选用 Bi-LSTM 网络来捕捉加速度信号的时间特征,以提高模型的性能。

Bi-LSTM 网络先将输入数据时序编码,然后将得到的时序特征向量送入输出层。为了更好地提取数据时序上的特征,Bi-LSTM 网络在 LSTM 的基础上结合了输入序列在前向和后向两个方向上的信息。Bi-LSTM 网络的运算图如图 3 所示,其中的横向黑色箭头表示信息的传输方向。该模块对波形的变化十分敏感且能有效捕捉相邻信息间的特征^[23]。

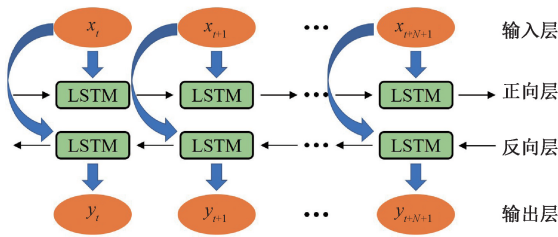


图 3 Bi-LSTM 网络运算图

Fig. 3 Operational graph of the Bi-LSTM network

在图 3 中,自左向右循环神经网络层更新公式为:

$$\vec{h}_t = H(W_{x\vec{h}} x_t + W_{h\vec{h}} \vec{h}_{t-1} + b_{\vec{h}}) \quad (2)$$

自右向左循环神经网络层的更新公式为:

$$\vec{h}_t = H(W_{x\vec{h}} x_t + W_{h\vec{h}} \vec{h}_{t-1} + b_{\vec{h}}) \quad (3)$$

前后两层循环神经网络层叠加后输出为:

$$y_t = W_{h_y} \vec{h}_t + W_{\vec{h}_y} \vec{h}_t + b_y \quad (4)$$

式中: t 表示输入的时间序列数据; $\vec{h}_t$ 表示在 t 时刻输出的隐藏层向量; x_t 表示 t 时刻的输入值; y_t 表示 t 时刻的输出值; $W_{x\vec{h}}$ 、 $W_{h\vec{h}}$ 表示输入 x_t 和隐藏层之间的权重矩阵; $W_{h\vec{h}}$ 、 $W_{\vec{h}h}$ 表示两个隐藏层之间的权重矩阵; W_{h_y} 、 $W_{\vec{h}_y}$ 表示隐藏层和输出层之间的权重矩阵; $b_{\vec{h}}$ 、 b_h 表示隐藏层的偏置向量; b_y 表示输出层的偏置向量; H 表示隐藏层的 ReLU 激活函数; 公式中的箭头代表信号方向。

1.3 特征融合

MCNN 对振动加速度信号进行空间上的特征提取,记为空间特征 S 。Bi-LSTM 网络对信号时域中数据依赖性的特征提取,记为时间特征 T 。最后,将空间特征 S 与时间特征 T 融合,获得一个新的融合特征 $[S, T]$ 。将 $[S, T]$ 送入分类层 (Softmax) 可以得到材料分类的预测结果,输出结果可以表示为:

$$\sigma = \text{Softmax}([S, T]) \in \mathbb{R}^n \quad (5)$$

式中:纹理有 n 类,由于本文使用的 LMT 触觉纹理数据库^[11]中有 9 类材料,故 $n=9$ 。

2 实验与讨论

本文通过消融实验和准确度对比实验对提出模型的有效性进行了检验。首先,介绍了所采用的数据集、数据扩增方法以及对训练集的准备。接着,通过网络架构优化选出最佳模块搭配方式。最后,给出了对比实验结果并分析了模型中不同组件的有效性和参数敏感度。

2.1 数据集

手动采集的振动加速度数据会随用户施加在纹理上的接触力和滑动速度的变化而变化。这种基于真实纹理交互过程建立的数据集包含了用户探索某一纹理表面时多种交互动作所对应的振动加速度数据,为纹理分类提供了多样化的特征,可使训练出的模型具备更好的泛化能力,并提高其鲁棒性。

在本文中,LMT 触觉纹理数据库中手动采集的振动加速度数据被用于训练提出的纹理分类模型。该数据库收集了用户与 69 个不同材料(如大理石、砖、木材等)纹理交互时的振动加速度数据,这些数据均由如图 4 所示的笔式装置以 10 kHz 的采样速率进行采集。存储在数据库中的振动加速度数据已经归一化,并通过 DFT321 算法将数据从三维转换为一维。对于每个纹理,已经记录了 10 个长度为 25 s 的振动加速度数据轨迹。



图 4 LMT 数据集的采集工具^[12]

Fig. 4 Collection tool for the LMT dataset^[12]

在真实的表面纹理探索场景中,用户手持工具在物体表面上滑动几百毫秒就足以判断该纹理的特征^[14]。因此,本文使用了 500 ms 的数据片段作为模型的输入。由于数据集是以 10 kHz 的速率进行采样,500 ms 的振动加速度数据包含有 5 000 个数据。

2.2 数据扩增

为了进一步提高模型对纹理分类的泛化能力,本文需要大量的振动加速度数据来对模型进行训练。由于 LMT 触觉纹理数据库中的数据量有限,本文使用滑动窗口数据采样来完成数据的扩增。基于滑动窗口的数据采样过程如图 5 所示。

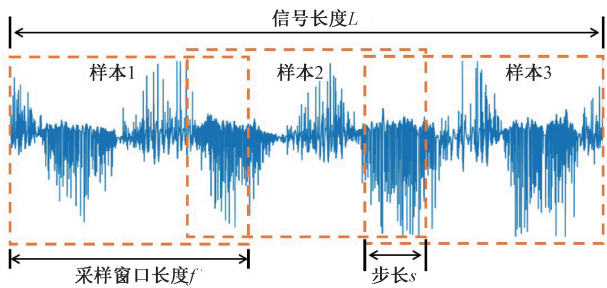


图 5 基于滑动窗口的数据采样

Fig. 5 Data sampling based on sliding window

设信号的总长度表示为 L , 采样窗口的长度表示为 f , 采样窗口移动步长表示为 s , 则数据数量 n 的计算公式为:

$$n = (L - f) / s \quad (6)$$

在原数据集中, 一个振动加速度数据轨迹的数据量为 250 000。设置滑动采样窗口长度 5 000, 步长 1 000, 则可对每个数据轨迹获得 245 组数据。10 个数据轨迹一共得到 2 450 组数据。与不采用滑动窗口采样的方法相比, 数据量扩增了 4.9 倍。

2.3 训练集准备

如图 6 所示, LMT 触觉纹理数据库收集的 69 个纹理表面按照物理属性可分为 9 类, 分别为序号 1~7 网格、序号 8~16 石头、序号 17~22 光泽表面、序号 23~31 木材、序号 32~35 橡胶表面、序号 36~43 纤维、序号 44~53 泡沫、序号 54~60 箔纸和序号 61~69 纺织品。每个纹理包含有用户手持刚性工具在真实材料表面上滑动 10 次(每次 25 s)的振动加速度数据。目前, 有以下 3 种方法来准备训练数据: 1) 将整个数据轨迹视为训练片段, 造成数据信息的冗余; 2) 使用预处理操作(如频谱图或 MFCC)将一维加速度信号转换为二维特征图, 忽略了数据时域特征, 特征提取不充分; 3) 从每个数据轨迹中采样一系列短片段, 可以获得大量的训练数据, 增强模型的泛化能力。

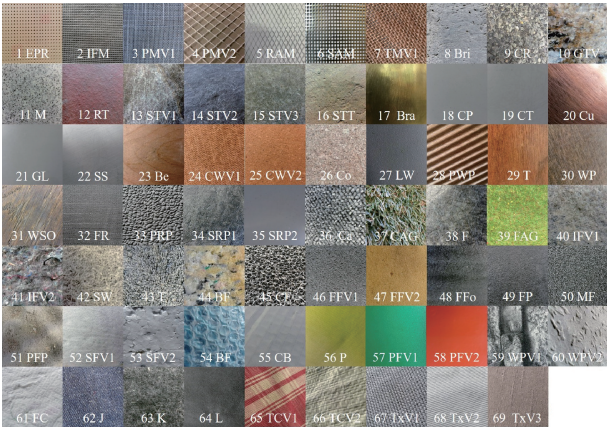


图 6 LMT 触觉纹理数据库中的材料分类图像^[12]

Fig. 6 Material classification images in the LMT tactile texture database^[12]

本文的实验采用方法 3) 准备训练数据, 使用 LMT 触觉纹理数据库中的振动加速度数据作为训练集, 并以时长为 0.5 s 的原始振动加速度数据片段作为网络的输入进行训练。在数据集中, 三轴的振动加速度数据已经利用 DFT321 算法组合成一维信号。该方法保留了三轴分量的时间和频谱特性, 并减少了由此产生的振动加速度信号对工具与纹理表面之间倾角的依赖性^[20]。

2.4 网络架构优化

1) LSTM 单元数量的选择

Bi-LSTM 网络用来提取加速度数据的时间依赖特征。该网络选取不同的 LSTM 单元数量时, 训练结果会有明显的差异。本文对包含有不同数量 LSTM 单元的 Bi-LSTM 模型进行了训练, 得到的结果如图 7 和 8 所示。图 7 表明, 随着 LSTM 单元数量的不断增加, 分类准确率会逐渐提升, 但增加到 128 个 LSTM 单元时, 准确率会低于 64 个 LSTM 单元的结果。图 8 显示了在不同数量的 LSTM 单元情况下模型损失值的差异。当模型的 LSTM 单元数量超过 64 个时, 训练得到的损失值会显著上升。这主要是因为模型的复杂度过高, 造成了过拟合问题。所以, 本文选择 64 个 LSTM 单元构建 Bi-LSTM 网络。

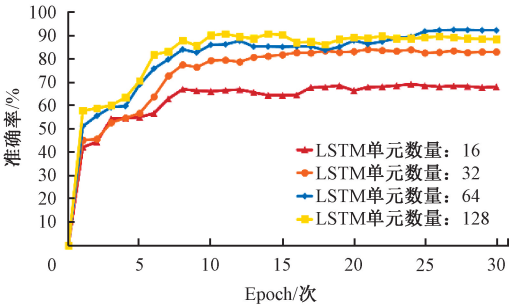


图 7 不同 LSTM 单元数量的模型准确率

Fig. 7 Model accuracy with different numbers of LSTM units

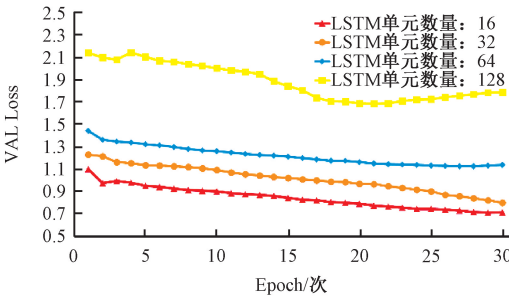


图 8 不同 LSTM 单元数量的模型验证损失值

Fig. 8 Model validation loss values for different numbers of LSTM units

2) MCNN 数量的选择

模型中的 MCNN 可以通过串联的方式进行堆叠, 以便通过不同大小的感受视野来挖掘重要的多尺度时间序

列特征。但是,过多的网络堆叠可能会在训练过程中产生过拟合现象。所以,针对不同数量的 MCNN 的模型架构,本文进行了性能对比预实验。在数据集上分别运行包含 1~6 个多尺度卷积子模块的网络架构,得到模型架构的搜索实验结果如图 9 所示。

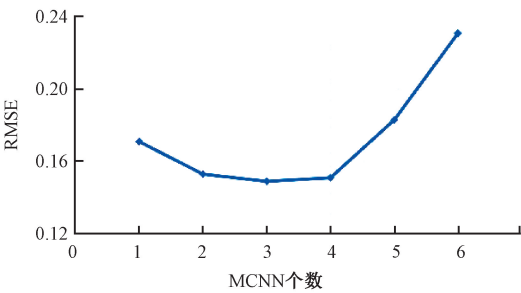


图9 不同 MCNN 个数的均方根误差

Fig. 9 Root mean square error of different MCNN numbers

从实验结果可以看出,当 MCNN 堆叠的个数 ≤ 3 时,模型的均方根误差随着网络个数的增加而降低,故增加网络的深度可以提高模型的性能。但当网络堆叠个数 >3 时,模型的均方根误差出现大幅度上升。而堆叠 3 个 MCNN 的架构取得了最低的均方根误差。因此,本文选择堆叠 3 个 MCNN 来组成纹理分类模型架构。

3) 模型架构的实现

经过对网络架构的优化,本文确定了由 3 个堆叠的 MCNN 与 Bi-LSTM 网络并联的纹理分类模型架构,如图 10 所示。

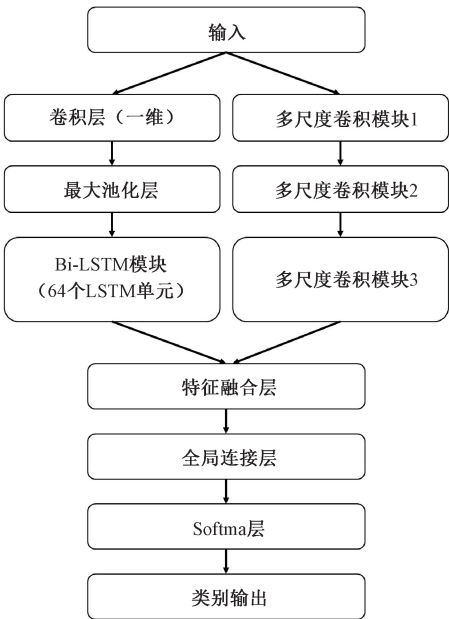


图10 优化后的模型架构图

Fig. 10 Optimized model architecture diagram

本模型训练策略如下:批处理大小设置为 256,初始学习率设置为 1×10^{-4} ,训练次数设置为 50 次,并采用 Adam 优化器和交叉熵损失函数。训练结果如图 11 所示,在接近 30 次时模型已经逐渐达到了最佳性能,最终的模型准确率稳定在 92% 左右。

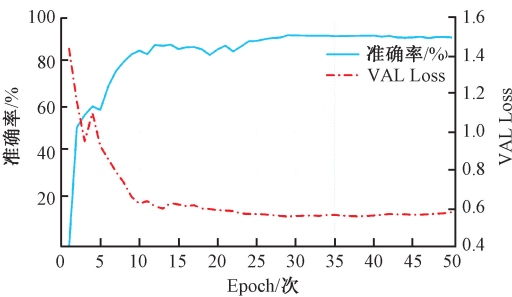


图11 模型的训练结果

Fig. 11 Training results of the model

2.5 消融实验

表 2 是不同神经网络方法的性能指标对比结果。由表 2 可知,在整体实验中,MCNN 和 Bi-LSTM 网络融合模型的性能明显优于单独的 LSTM 模型和 Bi-LSTM 模型,其分类准确率在测试集上达到 92.1%,远高于 Bi-LSTM 模型。

表 2 消融实验结果表

算法模型	实验结果		
	准确率	精确率	召回率
LSTM	71.4	69.7	85.5
Bi-LSTM	81.3	87.1	78.2
MCNN 和 Bi-LSTM 网络的融合	92.1	89.6	86.2

2.6 性能对比与分析

本文提出的方法与已有的 5 种纹理分类方法进行了性能比较。这 5 种方法分别为:

- 1) 基于集成学习的方法^[24]。该方法先从原始数据中进行手工特征提取,然后利用多聚类特征选择和拉普拉斯分值得特征选择算法来降低特征的维数。最后,提出了一种集成学习方法来综合不同特征选择方法的优点。该方法的有效性在 LMT 触觉纹理数据库上得到了验证。
- 2) 跨模态检索的方法^[25]。该方法先从原始数据中提取功率谱密度特征,然后应用主成分分析将功率谱密度特征降维。最后,这些功能被馈送到全连接的网络中。
- 3) 基于光谱图的方法^[26]。该方法先使用短时傅里叶变换将原始数据转换为频谱图,然后将其作为图像,利用二维 CNN 网络进行处理。

4) 基于一维 CNN 的方法^[14]。该方法直接将原始数据作为输入,并将来自 CNN 的特征展平作为纹理特征表示。

5) 基于一维 CNN-LSTM 的方法^[16]。该方法基于一维 CNN,但使用 LSTM 来整合时态信息。

如表 3 所示,本文所提出的模型对表面纹理的分类精确度较之前的方法有了明显的提升。而且本文提出的方

法是直接在振动加速度数据的基础上对纹理进行的分类,与基于集成学习的方法、跨模态检索的方法、基于光谱图的方法相比,一方面极大地降低了对数据来源的要求,另一方面在显著减少网络训练计算成本的情况下,实现了最佳的纹理分类性能。此外,从表 3 中也可以看出,本方法对原数据不需要做任何预处理,避免了手工特征提取的繁琐过程和数据二次变换带来的误差。

表 3 不同方法的分类性能对比

Table 3 Comparison of classification performance of different methods

方法	数据来源	数据预处理方式	分类模型	准确率/%
基于集成学习的方法 ^[24]	加速度数据、声音数据、纹理图像数据	手工特征提取	支持向量机	90.93
跨模态检索的方法 ^[25]	加速度数据、纹理图像数据	功率谱密度特征提取	DALNet	60.9
基于光谱图的方法 ^[26]	加速度数据、纹理图像数据	短时傅里叶变换	WCMALNet	84.3
基于 1D CNN 的方法 ^[14]	加速度数据	—	ACNN	81.8
基于 1D CNN-LSTM 的方法 ^[16]	加速度数据	—	CNN-LSTM	85.6
本文提出的方法	加速度数据	—	—	92.1

为了更好地理解提出模型的性能,图 12 中展示了模型的分类混淆矩阵图。混淆矩阵的每一行表示真实的纹理类别,每一列表示预测的纹理类别。图中的数值表示预测的纹理类别占该类别真实个数的比值,对角线比值越大,说明算法模型对该类别的分类准确率越高。

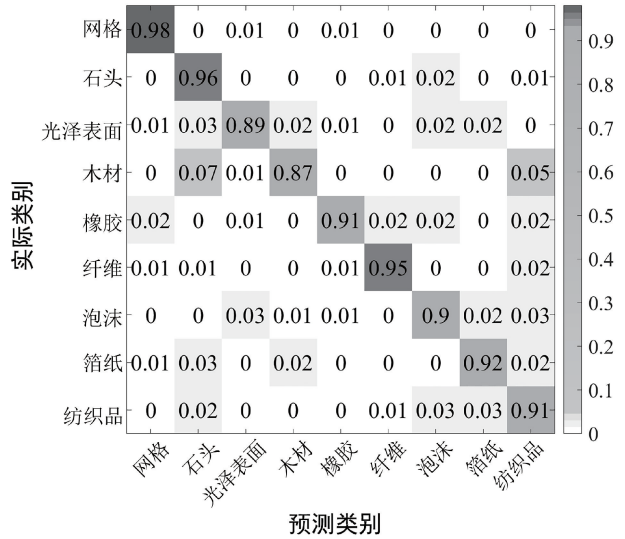


图 12 混淆矩阵

Fig. 12 Confusion matrix

从结果可以看到,模型总体纹理分类的平均精度达到 92.1%,其中本文的模型对网格类别的分类效果最佳,对木材的分类效果最差。原因可能是网格类纹

理有一定的规律性,采集到的数据也呈现一定的规则,故模型对其能够有较高的识别率。对木材类纹理的识别效果最差是由于模型部分地将其识别为石头或纺织品。这是因为木材类材料中有相当一部分纹理的振动加速度数据与石头或纺织品相似。最后,光泽表面和箔纸这类光滑表面纹理的振动加速度数据的时域特征与纹理粗糙的石头和纤维相比不够明显,识别准确率也相对较低。此外,基于纹理分类结果,整体 9 类纹理分类的准确率达到 92.1%。图 13 显示了部分材料的纹理图像和振动加速度数据。

2.7 泛化性能分析

为了进一步评估提出的纹理分类方法的泛化能力,本文将训练好的模型应用于另一个纹理触觉数据集——HaTT 数据集^[27]中与 LMT 数据集相同的纹理类型进行分类。如图 14 所示,本文按照 LMT 数据集中对纹理材质的归类方法,从 HaTT 数据集中选择了相应类别的 6 类不同纹理进行分类实验。每类纹理有 3 个不同的实例,每个实例采集了 10 s 的样本数据,并用 2.2 节描述的方式处理样本数据。然后将新数据集作为测试集。

如图 15 所示,提出的模型在 HaTT 数据集上的分类表现良好,总体平均准确率为 85.6%。除了对 HaTT 数据集中的光泽表面和木材的分类准确率稍低以外,模型对不同纹理类别的分类结果均较为准确。

以上实验结果表明,本文提出的模型不但对用于训练的 LMT 数据集具有较高的纹理分类准确度,还对具有

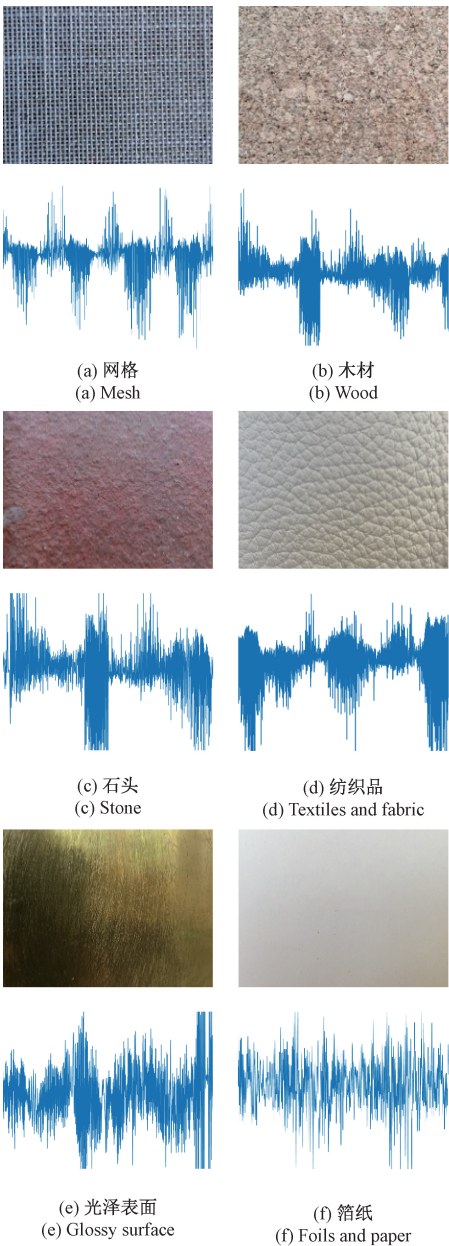


图 13 不同材料的纹理图像和振动加速度数据
Fig. 13 Texture images and vibration acceleration data of different materials

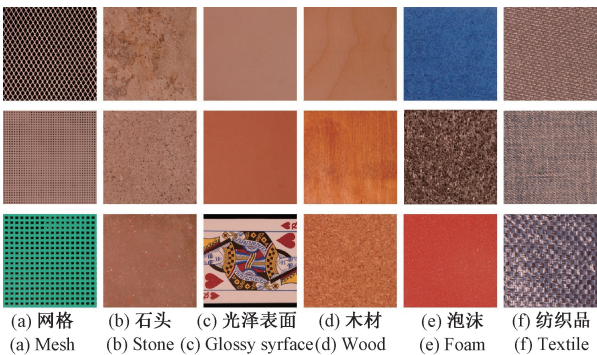


图 14 从 HaTT 数据集中选取的 6 类纹理
Fig. 14 Six types of textures selected from the HaTT dataset

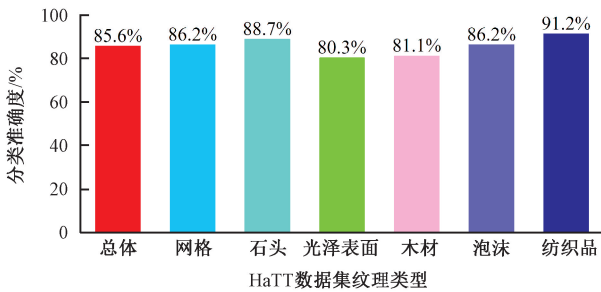


图 15 模型在 HaTT 数据集上对纹理的分类准确率
Fig. 15 The accuracy of texture classification by the model on the HaTT dataset

类似纹理的其他数据集具有较好的泛化性能。因此,本文提出的模型具有良好的应用前景。

3 结 论

为了探究触觉纹理数据序列间的时间相关性以及不同时间尺度信息对纹理分类效果的影响,进一步提高分类准确率,本文提出一种 MCNN 和 Bi-LSTM 网络相结合的融合模型用于材料表面纹理分类。与其他方法在 LMT 触觉纹理数据库上进行了分类效果的比较,结果表明,相比其他 5 种已有的纹理分类方法,本模型的性能更优,从整体上提高了分类准确率。经消融实验与对比实验可知,MCNN 具有多个不同大小的特征感受视野,可以提取振动加速度数据在不同尺度上的特征,得到更充足的特征表示。Bi-LSTM 网络具备双向提取时间序列特征的能力,使得模型可以更加有效地学习时间序列数据中的相关性信息,进一步提高了纹理识别的准确率。该网络将两个支路以并行的方式组合,使模型自动学习到振动加速度信号的空间与时间特征。

该模型分类效果与以往的研究工作相比,有了较大程度上的提升,但仍有很大的改进空间,需要开展进一步的研究。未来的研究工作可以从以下两个方面开展。首先,仅从触觉数据的时域方面对纹理分类开展研究会忽略对触觉数据同样重要的频域特征。在后续的研究中可以融合对触觉数据的频域分析来完成纹理分类的深度学习建模工作,进一步提高分类的准确度。其次,本文目前只针对触觉振动加速度数据开展了纹理分类研究,但随着集成多种传感器的触觉装置的发展,多触觉变量融合的分类方式能带来更多的特征信息,从而更好地完成分类任务。在以后的工作中,将在此模型的基础上深入研究多触觉变量分类问题的模型架构。

参考文献

[1] ECK J, KAAS A L, GOEBEL R. Crossmodal interactions of haptic and visual texture information in

- early sensory cortex [J]. *Neuroimage*, 2013, 75: 123-135.
- [2] 刘丽, 赵凌君, 郭承玉, 等. 图像纹理分类方法研究进展和展望[J]. *自动化学报*, 2018, 44(4): 584-607.
- LIU L, ZHAO L J, GUO CH Y, et al. Texture classification: State-of-the-art methods and prospects [J]. *Acta Automatica Sinica*, 2018, 44(4): 584-607.
- [3] WEI J, CUI S, HU J, et al. Multimodal unknown surface material classification and its application to physical reasoning [J]. *IEEE Transactions on Industrial Informatics*, 2021, 18(7): 4406-4416.
- [4] 余玉卿, 宋爱国, 陈大鹏, 等. 用于触摸屏图像感知的指端力触觉再现系统[J]. *仪器仪表学报*, 2017, 38(6): 1523-1530.
- YU Y Q, SONG AI G, CHEN D P, et al. Fingertip haptic rendering system for touch screen image perception [J]. *Chinese Journal of Scientific Instrument*, 2017, 38(6): 1523-1530.
- [5] 陈大鹏, 高亚洲, 宋爱国, 等. 基于触摸屏交互的指套式盲文再现系统[J]. *仪器仪表学报*, 2022, 43(5): 199-208.
- CHEN D P, GAO Y ZH, SONG AI G, et al. Fingertip braille display system based on touch screen interaction [J]. *Chinese Journal of Scientific Instrument*, 2022, 43(5): 199-208.
- [6] LEDERMAN S J, KLATZKY R L. Haptic perception: A tutorial [J]. *Attention, Perception, & Psychophysics*, 2009, 71(7): 1439-1459.
- [7] OKAMOTO S, NAGANO H, YAMADA Y. Psychophysical dimensions of tactile perception of textures [J]. *IEEE Transactions on Haptics*, 2013, 6(1): 81-93.
- [8] STRESE M, LEE J Y, SCHUWERK C, et al. A haptic texture database for tool-mediated texture recognition and classification [C]. *IEEE International Symposium on Haptic, Audio and Visual Environments and Games (HAVE)*, 2014: 118-123.
- [9] JAMALI N, SAMMUT C. Majority voting: Material classification by tactile sensing using surface texture [J]. *IEEE Transactions on Robotics*, 2011, 27(3): 508-521.
- [10] CULBERTSON H, UNWIN J, KUCHENBECKER K J. Modeling and rendering realistic textures from unconstrained tool-surface interactions [J]. *IEEE Transactions on Haptics*, 2014, 7(3): 381-393.
- [11] STRESE M, SCHUWERK C, STEINBACH E. Surface classification using acceleration signals recorded during human freehand movement [C]. *IEEE World Haptics Conference (WHC)*, 2015: 214-219.
- [12] FISHEL J A, LOEB G E. Bayesian exploration for intelligent identification of textures [J]. *Frontiers in Neurobotics*, 2012, 6-4: 1-20.
- [13] NIE D, LIU J, SUN X. Influence of surface tactile data quantity on material classification in unstructured environments [J]. *IEEE Transactions on Instrumentation and Measurement*, 2021, 70: 1-12.
- [14] JI M, FANG L, ZHENG H, et al. Preprocessing-free surface material classification using convolutional neural networks pretrained by sparse autoencoder [C]. *IEEE 25th International Workshop on Machine Learning for Signal Processing (MLSP)*, 2015: 1-6.
- [15] ORII H, TSUJI S, KOUUDA T, et al. Recurrent neural network for tactile texture recognition using pressure and 6-axis acceleration sensor data [C]. *Tencon 2017-2017 IEEE Region 10 Conference*, 2017: 2012-2016.
- [16] BEDNAREK J, BEDNAREK M, KICKI P, et al. Robotic touch: Classification of materials for manipulation and walking [C]. *IEEE 2nd International Conference on Soft Robotics (RoboSoft)*, 2019: 527-533.
- [17] 惠文珊, 李会军, 陈萌, 等. 基于 CNN-LSTM 的机器人触觉识别与自适应抓取控制 [J]. *仪器仪表学报*, 2019, 40(1): 211-218.
- HUI W SH, LI H J, CHEN M, et al. Robotic tactile recognition and adaptive grasping control based on CNN-LSTM [J]. *Chinese Journal of Scientific Instrument*, 2019, 40(1): 211-218.
- [18] JOOLLEE J B, UDDIN M A, JEON S. Deep multi-model fusion network based real object tactile understanding from haptic data [J]. *Applied Intelligence*, 2022: 1-16.
- [19] METZGER A, TOSCANI M. Unsupervised learning of haptic material properties [J]. *Cold Spring Harbor Laboratory*, 2021, DOI:10.1101/2021.02.25.432896.
- [20] LANDIN N, ROMANO J M, MCMAHAN W, et al. Dimensional reduction of high-frequency accelerations for haptic rendering [C]. *International Conference on Human Haptic Sensing and Touch Enabled Computer Applications*, 2010: 79-86.
- [21] CUI Z, CHEN W, CHEN Y. Multi-scale convolutional neural networks for time series classification [J]. *ArXiv Preprint*, 2016, ArXiv:1603.06995.
- [22] RAHUL J, SHARMA L D. Automatic cardiac arrhythmia classification based on hybrid 1-D CNN and Bi-LSTM model [J]. *Biocybernetics and Biomedical Engineering*, 2022, 42(1): 312-324.
- [23] 李洋, 董红斌. 基于 CNN 和 BiLSTM 网络特征融合的

文本情感分析[J]. 计算机应用, 2018, 38(11): 3075-3080.

LI Y, DONG H B. Text sentiment analysis based on feature fusion of convolution neural network and bidirectional long short-term memory network [J]. Journal of Computer Applications, 2018, 38(11): 3075-3080.

- [24] LIU X, WU H, FANG S, et al. Multimodal surface material classification based on ensemble learning with optimized features [C]. IEEE International Conference on E-health Networking, Application & Services, 2021: 1-6.
- [25] ZHENG W, LIU H, WANG B, et al. Cross-modal surface material retrieval using discriminant adversarial learning [J]. IEEE Transactions on Industrial Informatics, 2019, 15(9): 4978-4987.
- [26] ZHENG W, LIU H, WANG B, et al. Cross-modal material perception for novel objects: A deep adversarial learning method [J]. IEEE Transactions on Automation Science and Engineering, 2019, 17(2): 697-707.
- [27] CULBERTSON H, DELGADO J J L, KUCHENBECKER K J. One hundred data-driven haptic texture models and open-source methods for rendering on 3D objects [C]. IEEE Haptics Symposium (HAPTICS), 2014: 319-325.

作者简介



陈大鹏, 2011 年于安徽理工大学获得学士学位, 2019 年于东南大学获得博士学位, 现为南京信息工程大学副教授、硕士生导师, 主要研究方向为力触觉再现、人工智能、人机交互。

E-mail: dpchen@nuist.edu.cn

Chen Dapeng received his B.Sc. degree from Anhui University of Science and Technology in 2011, and received his

Ph. D. degree from Southeast University in 2019. He is currently an associate professor and a master advisor at Nanjing University of Information Science & Technology. His main research interests include haptic display, artificial intelligence, and human-computer interaction.



刘佳(通信作者), 2008 年于东南大学获得博士学位, 现为南京信息工程大学自动化学院教授、硕士生导师, 主要研究方向为力触觉再现、虚拟/增强现实、人机交互。

E-mail: liujia@nuist.edu.cn

Liu Jia (Corresponding author) received her Ph. D. degree from Southeast University in 2008. She is currently a professor and a master advisor in the School of Automation at Nanjing University of Information Science & Technology. Her main research interests include haptic display, virtual/augmented reality, and human-computer interaction.



宋爱国, 分别在 1990 年和 1993 年于南京航空航天大学获得学士学位和硕士学位, 1996 年于东南大学获得博士学位, 现为东南大学仪器科学与工程学院教授、博士生导师, 主要研究方向为力触觉再现、康复工程、机器人遥操作技术。

E-mail: a.g.song@seu.edu.cn

Song Aiguo received his B.Sc. and M.Sc. degrees both from the Nanjing University of Aeronautics and Astronautics in 1990 and 1993, respectively, and received his Ph. D. degree from Southeast University in 1996. He is currently a professor and a Ph. D. advisor in the School of Instrument Science and Engineering at Southeast University. His main research interests include haptic display, rehabilitation engineering, and robotic teleoperation technology.