

基于新型联合感知矩阵的压缩学习目标识别技术

李 峰¹, 詹邦成^{1,2}, 辛 蕾¹, 刘 洋¹, 刘志佳³, 肖化超⁴

(1. 钱学森空间技术实验室, 北京 100094; 2. 河南大学河南省大数据分析与管理重点实验室, 河南开封 475004;
3. 航天东方红卫星有限公司, 北京 100094; 4. 中国空间技术研究院西安分院, 陕西西安 710100)

摘要: 目标识别正逐渐成为自动化领域中提供准确目标类别信息的一项重要技术, 并且当前大多数目标识别方法都是基于深度学习框架实现. 通常, 深度学习框架的输入数据均为原始图像数据, 而在实际应用中, 探测器获取原始图像数据并作为深度学习框架的输入进而实现目标识别的方式并非高效的, 数据获取并识别的过程包含了大量的冗余信息, 降低了识别效率. 在本文中, 通过深度学习与压缩感知技术的结合, 提出了一种基于联合感知矩阵的压缩学习目标识别技术 (Target recognition technology based on a new joint sensing matrix for compressed learning, TRNPCL), 使得探测器可快速生成目标图像多维压缩数据, 且压缩数据可直接作为深度学习目标识别框架的输入数据, 而无需再进行解压缩步骤. 该方法不仅大大减小了深度学习框架的数据输入量, 在与同等压缩比下的单空间域数据压缩学习方式相比较, 还保持了较高的识别准确率. 在未来, 该方法有望成为一种更有效、更灵活的目标识别方法, 并特别适用于指纹识别、人脸识别等应用领域.

关键词: 压缩感知; 深度学习; 目标识别; 压缩编码; 图像处理

中图分类号: TP391.4

文献标识码: A

文章编号: 0372-2112(2021)11-2108-09

电子学报 URL: <http://www.ejournal.org.cn>

DOI: 10.12263/DZXB.20200852

Target Recognition Technology Based on a New Joint Sensing Matrix for Compressed Learning

LI Feng¹, ZHAN Bang-cheng^{1,2}, XIN Lei¹, LIU Yang¹, LIU Zhi-jia³, XIAO Hua-chao⁴

(1. Qian Xuesen Laboratory of Space Technology, Beijing 100094, China;

2. Henan Key Laboratory of Big Data Analysis and Processing, Henan University, Kaifeng, Henan 475004, China;

3. DFH Satellite Co., Ltd., Beijing 100094, China;

4. China Academy of Space Technology (Xi'an), Xi'an, Shaanxi 710100, China)

Abstract: Target recognition is gradually becoming an important technology to provide accurate target category information in the field of automation, and most current target recognition methods are based on machine learning frameworks. Generally, the input data of the machine learning framework is raw image data, but in practical applications, the way that the detector obtains the raw image data and uses it as the input of the deep learning framework to achieve target recognition is low efficient, contains a lot of redundant information, which reduces the recognition efficiency. In this paper, by combining machine learning and compressive sensing technology, a new target recognition method named as target recognition technology based on a new joint sensing matrix for compressed learning (TRNPCL) is proposed. Through the proposed method, the detector can quickly generate multi-dimensional compressed data of the target image, and the compressed data can be directly used as the input data of the deep learning target recognition framework without further decompression steps, which not only greatly reduces the amount of data input to the machine learning framework, but also maintains a higher recognition accuracy compared with the single-space domain data compression learning method under the same compression ratio. In the future, this method is expected to become a more effective and flexible target recognition method, and especially suitable for fingerprint recognition, face recognition and other application fields.

Key words: compressed sensing; deep learning; target recognition; compression coding; image processing

1 引言

图像有着强大的信息搭载能力,可详细记录人类经济社会发展过程中所产生的海量可视化信息,随着科技的发展,基于图像开展的目标识别技术研究已经成为计算机视觉研究领域中的重要组成部分,其智能化的处理方式,对人类高科技研究领域的发展也产生了深远的影响,该方法不仅成为军事目标快速识别的关键技术,而且是远程无人机、机器人等民用智能设备系统的核心部分,还在航天、医学等领域应用广泛。

近年来,随着图像数据量的不断增大,目标类型更加复杂,纹理信息更加丰富,通过在图像中找到对应匹配的位置,计算相似度判定是否是相应目标的传统目标识别方法,已远远不能满足现在的需求,具备高效率、高准确率的基于深度学习的图像目标识别技术逐渐崭露头角。

基于深度学习的图像目标识别技术在近年来取得了多项成果,如 AlexNet^[1], GoogleNet^[2] 等经典目标识别神经网络模型的出现,以及当今前沿的目标检测算法 (R-CNN^[3], SPP (Spatial Pyramid Pooling) NET^[4], Faster R-CNN^[5], YOLO^[6] 等) 中目标识别模块的成功应用,都见证了目标识别技术的快速发展,综合其发展趋势可概括为:模型层次不断加深,结构更加复杂,训练数据不断增加,使得基于深度学习的目标识别技术在实际应用过程中也面临着一些问题,如基于深度学习的人脸识别系统在运行过程中识别数据量大,数据分配效率低,如何加快其处理速率是有待解决的问题之一^[7];在指纹识别领域,基于神经网络的指纹识别方式存在的计算数据量大的问题,大大限制了基于深度学习的指纹识别技术的推广与应用,使得目前常用的指纹识别技术多是基于提取关键点进行匹配的传统方式实现^[8]。如何提升目标识别神经网络的效率,并能够保持其较高的识别准确度效果,是现今目标识别技术在持续发展过程中有待解决的问题。

合理的数据压缩设计不仅能够高效利用成像设备的存储容量,而且将大大减少深度学习神经网络的数据输入量,降低神经网络的计算量,提升其效率,若还能使其原始目标数据特征得到最大限度的保留,维持其较高的识别准确度,则是一举多得的方法。

传统的数据压缩普遍采用基于数字余弦变换、小波变换的压缩方法^[9],但缺点也显而易见:效率低、复杂度高、功耗高、可移植性差等,而目前使用的图像压缩算法 (JPEG2000 或 SPIHT^[10]) 较为复杂,能耗都比较高,并且需要额外的压缩芯片实现数据图像压缩,使得在探测器在进行数据获取过程中综合效率不高。近年来,压缩感知理论 (Compressive Sensing, CS)^[11] 引起了广泛关注,压缩感知的理论基础是由 Emmanuel Candès, Ter-

ence Tao 和 David Donoho 建立,将其“采样”与“压缩”合二为一的方式,可实现对原始图像信号进行简单的基于感知矩阵进行欠奈奎斯特 (Shannon-Nyquist) 采样来获得压缩数据,即探测器感光元件读出的信号就是已经压缩好的信号,无需多余的图像压缩模块进行压缩。这种方式不仅为数据获取提供了更高的效率,且对硬件设计更加友好,若将压缩感知应用于硬件图像获取-算法目标识别一体化目标识别技术中,将大大提升识别效率。

压缩感知理论已应用于许多领域,最成功的例子之一是美国莱斯大学开发的单像素数码相机。单像素相机^[12,13]将图像采集和图像压缩集成在一起,即在采样过程中实现图像压缩,大大降低光电探测器设计的复杂度和成本。

在过去十年中,压缩感知多被应用于信号转换,并且大量的工作都致力于更好的重建信号,其中数据压缩过程获取的观测值数据结果也一直被作为图像重建的基础^[14,15],但使用相同压缩步骤获取的观测值数据,其实也保留了图像数据在压缩域的潜在判别特征^[16,17],基于压缩感知获取的观测值进行进一步探索同样也具备很大的意义。针对这一思路,本文结合压缩感知图像压缩编码与深度学习技术,提出了一种基于联合感知矩阵的压缩学习目标识别技术,以实现快速高效的目标识别应用。

2 方法及步骤

基于新型联合感知矩阵的压缩学习目标识别技术首先在探测器内部基于构建的新型联合感知矩阵 (Joint-Perception-Matrix, JPM) 高效获取具备多域压缩特征信息的多维压缩数据,再结合目标识别卷积神经网络实现目标识别,使其不仅在空间压缩域,还在压缩感知域同步进行特征提取,通过多域特征学习的方式,弥补了数据压缩带来的特征信息丢失的问题,从而在目标识别过程中不仅提升了识别效率,而且保持了较高的识别准确率。基于新型联合感知矩阵的压缩学习目标识别技术可划分为基于新型联合感知矩阵的多维压缩数据获取与基于目标分类识别卷积神经网络的图像目标识别方法两个模块,其框架如图 1 所示。

2.1 多维压缩数据获取方法

多维压缩数据的获取可基于新型联合感知矩阵实现。该新型联合感知矩阵 $\Phi_{\zeta} = \mathbf{R}^{M \times N}$ 由两个子矩阵构成,即 $\Phi_{\zeta} = \begin{bmatrix} \Phi_{\zeta D} \\ \Phi_{\zeta N} \end{bmatrix}$ 。其中, $\Phi_{\zeta D} \in \mathbf{R}_{\zeta}^{M_2}$ 为缩略图下采样矩阵,该下采样矩阵获取的下采样图可以很大程度的保留图像在空间域的边缘轮廓特征信息; $\Phi_{\zeta N} \in \mathbf{R}_{\zeta}^{M_1}$ 为传统压缩感知的观测值数据,且 ($M = M_1 + M_2$), 可认为通

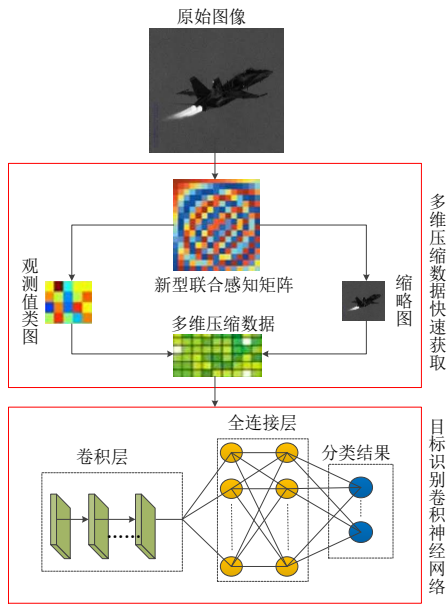


图1 基于联合感知矩阵的压缩学习目标识别技术框架

过变换后保留了图像在多压缩域的特征。

在硬件设计上,将 Φ_{cD} 与 Φ_{cN} 的获取完全基于探测器内部实现. 将感光元件获取的数据变换采样后直接可作为 Φ_{cN} ; 在探测器内部使 Binning 功能可实现原始图像数据的下采样, 进而实现 Φ_{cD} 的获取, 其流程如图2所示.

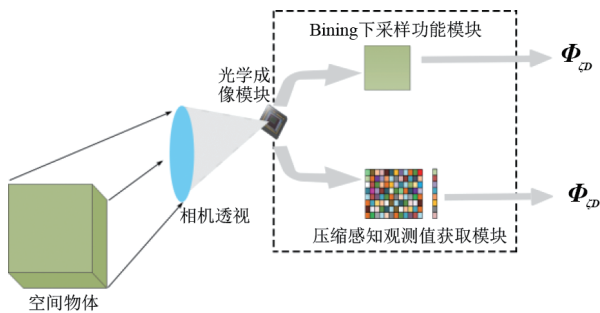


图2 探测器内部多类型数据获取方式

在生成观测值数据的变换方式上,常进行 Noiselet 变换^[18],但 Noiselet 变换矩阵中存在虚部,在工程应用中虚部的实现比较困难,因此本文采用 Real-Noiselet 变换,一种不包含虚部,只包含实数的数据变换方式.

对于一个长度为 n 的向量 $\mathbf{x} = [x_1, x_2, \dots, x_{n-1}, x_n]$,使用图3所介绍的 Real-Noiselet 变换方法对其进行 Real-Noiselet 变换,最终可以得到变换后的向量 $\mathbf{f} = [f_1, f_2, \dots, f_{n-1}, f_n]$.

整个 Real-Noiselet 变换过程包含 m 步, $m = \log_2 n$, $\mathbf{x} = [x_1, x_2, \dots, x_{n-1}, x_n]$ 经过第一步变换得到 $[x_1^{(1)}, x_2^{(1)}, \dots, x_{n-1}^{(1)}, x_n^{(1)}]$,具体变换如式(1)所示:

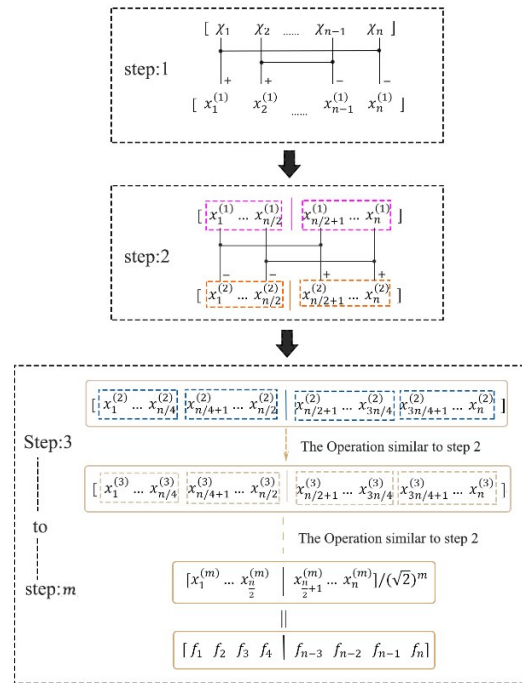


图3 Real-Noiselet 变换方式

$$\begin{cases} x_i^{(1)} = x_i + x_{n-i+1}, & i = 1, \dots, \frac{n}{2} \\ x_{n-i+1}^{(1)} = x_i - x_{n-i+1} \end{cases} \quad (1)$$

以中间为对称轴,第一个元素 x_1 与最后一个元素 x_n 相加($x_1 + x_n$)作为第一步变换结果的第一个元素 $x_1^{(1)}$,第一个元素与最后一个元素相减($x_1 - x_n$)作为第一步变换结果的最后一个元素 $x_n^{(1)}$.以同样的方法处理剩下的 $(n - 2)$ 个元素,得到的 $[x_1^{(1)}, x_2^{(1)}, \dots, x_{n-1}^{(1)}, x_n^{(1)}]$ 中的每个元素都包含 x 中 2 个元素的信息.

对于第二步变换,将 $[x_1^{(1)}, x_2^{(1)}, \dots, x_{n-1}^{(1)}, x_n^{(1)}]$ 平均分成两部分,每一部分看作为一个窗,每个窗包含 $n/2$ 个元素,然后将两个窗看为一组,第一个窗中的第一个元素与第二个窗中的第一个元素相减的结果 $x_1^{(1)} - x_{\frac{n}{2}+1}^{(1)}$ 放入第一个窗的第一个位置,即得到了 $x_1^{(2)}$.第一个窗中的第一个元素与第二个窗中的第一个元素相加的结果 $x_1^{(1)} + x_{\frac{n}{2}+1}^{(1)}$ 放入第二个窗的第一个位置 $x_{\frac{n}{2}+1}^{(2)}$,然后分别对两个窗内的第二个元素做相同的操作,得到 $x_2^{(2)}$ 与 $x_{\frac{n}{2}+2}^{(2)}$,依次对剩余的元素进行此操作,得到 $[x_1^{(2)}, x_2^{(2)}, \dots, x_{n-1}^{(2)}, x_n^{(2)}]$,如式(2)所示,得到的结果中的每一个元素都包含 x 中 4 个元素的信息.

$$\begin{cases} x_i^{(2)} = x_i^{(1)} - x_{\frac{n}{2}+i}^{(1)}, & i = 1, \dots, \frac{n}{2} \\ x_{\frac{n}{2}+i}^{(2)} = x_i^{(1)} + x_{\frac{n}{2}+i}^{(1)} \end{cases} \quad (2)$$

之后的第三步到第 m 步变换都与第二步类似,只

是窗的大小越来越小,每一步包含的元素个数为上一步窗包含个数的 1/2,直到最终每个窗只包含 2 个元素,在每一步变换过程中,都以中间为对称轴,对称的两个窗作为一组进行操作. 经过 m 步变换,最终得到的 $f = [f_1, f_2, \dots, f_{n-1}, f_n]$ 中每一个元素都包含 x 中全部元素的信息,图 4 给出了一幅航空影像的 Real-Noiselet 变换流程.

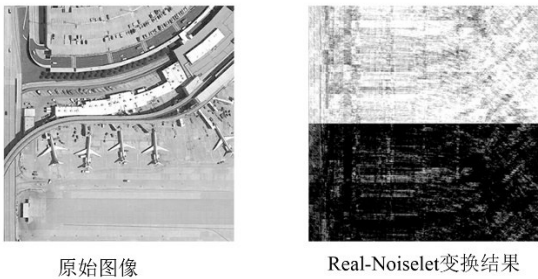


图 4 原始图像及其经过 Real-Noiselet 变换图示

经过 Real-Noiselet 变换,图像原始信息基本被打散到了压缩域,变换得到的图像整体特征较为弥散,但是保留了原始图像的全部信息,即通过这种变换方式将图像的特征分布在压缩域. 然后对得到的 Real-Noiselet 最终变换结果进行采样压缩,为了最大程度的保留图像信息的规律性,本文并未对观测数据进行随机采样,而是通过定点下采样方式对观测数据进行采样压缩. 定点下采样方式是基于本文的 Real-Noiselet 变换数据的特殊性进行的数据压缩方式,其与传统的下采样方式极其类似,只是在选择采样点时通过固定采样点的方式进行,例如进行的是 4 倍压缩,则压缩采样方式如图 5 所示,对于每个 2×2 pixel 滑窗中仅保留左上角的像素点作为采样点,最后生成新的 Real-Noiselet 变换压缩数据.

通过上述定点采样方法获取的 Real-Noiselet 压缩变换数据就是最终的观测值压缩数据. 若进行 τ 倍压缩,则自上而下,自左向右依次保留 $\sqrt{\tau} \times \sqrt{\tau}$ 窗口大小

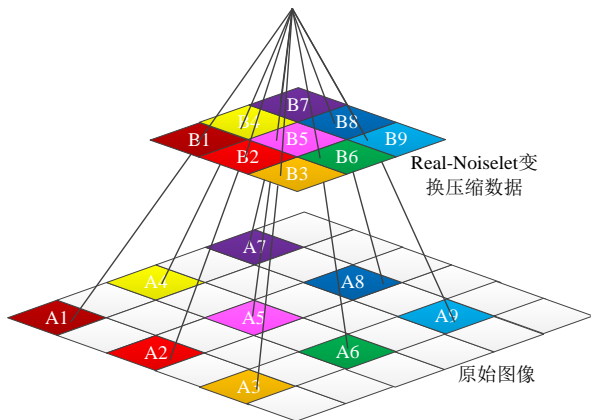


图 5 对观测值实现 4 倍压缩采样流程图示

的右上角像素作为采样点,生成观测值压缩数据. 原始图像再通过探测器内部 Binning 采样模块进行传统下采样可获取缩略图数据. 基于该新型联合感知矩阵获取的观测值压缩数据可以作为类图像数据,与下采样缩略图数据可组成多维度的原始图像压缩描述数据,其大小将小于原始数据,且具备多维度的压缩数据特征,可代替原始数据作为目标识别神经网络的输入数据.

2.1 基于压缩学习的目标识别神经网络构建

多维压缩数据具备广泛的数据特征信息,其包含的目标观测值类图 ω 包含了目标在图像压缩域的特征分布信息,缩略图像 Ψ 包含了目标在空间域的特征分布信息,生成的多维压缩数据类图 $\alpha = \begin{bmatrix} \Psi \\ \omega \end{bmatrix}$ 可直接作为目标识别神经网络的输入数据进行识别,本文中目标识别卷积神经网络的构建方案如图 6 所示.

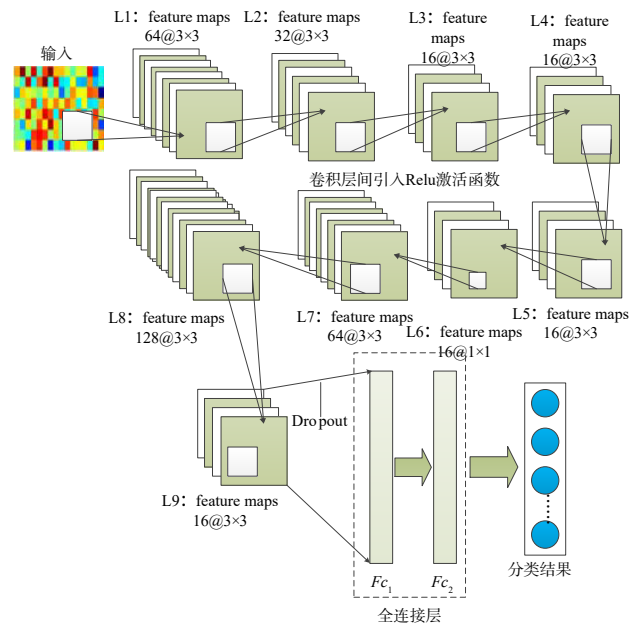


图 6 目标识别神经网络框架

其识别原理是将 N 组多维压缩数据 $\{\alpha_1, \alpha_2, \dots, \alpha_N\}$ 通过目标识别神经网络进行训练,使得训练获得的模型 β 既学习到目标的空间域特征也学习到目标在压缩域的信息分布特征,以此判定目标的类别.

该目标识别神经网络的卷积层^[19,20]设置为 9 层 $\{L_1, L_2, L_3, \dots, L_9\}$,卷积核的大小及个数见表 1.

为了增加卷积层之间的非线性关系,引入了 ReLU^[21]作为激活函数,并且在神经网络中多层卷积层后引入 dropout^[22]操作,dropout 操作使得在训练模型时,随机让网络的某些节点不工作(输出置零),也不更新权重(但会保存下来,只是本次训练不参与 BP (back-propagating)^[23]传播,其它过程不变). Dropout 操

表1 目标分类神经网络卷积层构造表

卷积层	卷积核数量	卷积核大小
L_1	64	(3×3)
L_2	32	(3×3)
L_3	16	(3×3)
L_4	16	(3×3)
L_5	16	(3×3)
L_6	16	(1×1)
L_7	64	(3×3)
L_8	128	(3×3)
L_9	16	(3×3)

作可以比较有效地减少过拟合的发生,一定程度上达到了正则化的效果,而后接入双层全连接层 $\{F_{c1}, F_{c2}\}$,每个全连接层均包含128个神经元。

最后在输出端使用SOFTMAX^[24]激励函数,用于离散化概率分布,将结果映射到 $[0, 1]$ 中,实现输出结果的多分类识别。

3 实验验证

3.1 基于新型联合感知矩阵的数据快速获取实验验证

实验部分对传统原始图像数据获取方法与基于新型联合感知矩阵的多维压缩数据获取方法进行了A、B两组实验对比验证,并对其获取速率 ϑ 和数据所占存储空间大小 ε 两方面指标数据进行统计分析。其获取速率 ϑ 指的平均每秒生成的图像数据量(张/s); ε 为获取的 n 张影像所需的存储空间容量;压缩比 Q 的计算方式如式(3)所示:

$$Q = \frac{Z}{\mathbb{R}} \quad (3)$$

其中 Z 为原始图像数据大小, \mathbb{R} 为 Z 压缩后的图像数据大小。

(1) 实验A

实验A部分,分别获取了图像大小为 256×256 pixel的原始图像数据 \mathcal{A}_1 ,以及基于新型联合感知矩阵的4倍压缩比多维压缩数据 \mathcal{B}_1 ,根据实验结果,对不同方法的数据获取速率与数据占用存储空间大小进行了统计,相应数据如表2所示。

(2) 实验B

实验B部分,分别获取了大小为 512×512 pixel的原

表2 4倍压缩数据生成实验数据对比表

	压缩比 Q_1	数据生成速率 ϑ_1 (张/s)	2000张数据存储空间 大小 ε_1 (MB)
\mathcal{A}_1		46	27.0
\mathcal{B}_1	4	60	12.4

始图像数据 \mathcal{A}_2 ,以及基于新型联合感知矩阵的8倍压缩比多维压缩数据 \mathcal{B}_2 ,根据实验结果,对数据获取速率与数据的占用存储空间大小进行了统计,相应数据如表3所示。

表3 8倍压缩数据生成实验数据对比表

	压缩比 Q_2	数据生成速率 ϑ_2 (张/s)	2000张数据存储空间 大小 ε_2 (MB)
\mathcal{A}_2		19	79.9
\mathcal{B}_2	8	64	16.8

对实验A、B的结果数据进行分析可知,基于新型联合感知矩阵的数据获取方法在数据生成速率方面更加高效,并且在对于同样数量的数据图像,多维压缩数据比原始图像数据占用更小的存储空间。即本文提出的基于新型联合感知矩阵的数据获取方法与传统数据获取方法相比在数据生成效率与存储空间占用率方面具备优势。

3.2 基于联合感知矩阵的压缩学习目标识别技术实验验证

实验验证部分选用统一的实验环境,实验环境配置如表4所示。

表4 实验环境配置

硬件环境	GeForce GTX 1660 Ti graphics card; CPU: Intel® Core™ i7-9750H; Memory: 16GB.
软件环境	Pycharm-community-2019. 3. 3; Anaconda3-5. 3. 1-Windows-x86_64; Cuda version: 10. 1; CudaNN version: 7. 6. 5; Tensorflow-GPU 2. 1. 0.

实验中选取Cifar-10^[25]数据集以及MNIST^[26]数据集作为实验数据,实验效果评价指标如下:

设测试图像总个数为 Ω ,其中正确识别个数为 β ,错误识别个数为 γ ,训练时长为 δ ;对于同一个类别目标,正确识别为该类型目标的数量为 t_1 ,错误识别为该类型目标的数量为 t_2 ,则评价指标定义如表5所示。

表5 实验评价指标

训练效率 δ	δ 的大小可用作测量标准。在相同条件下, δ 越大,计算量越大,效率越低; δ 越小,计算量越小,效率越高。
精准度 ϖ	$\varpi = \frac{\beta}{\Omega}$
模型识别效率 ω	ω 的大小可用作衡量标准。在相同条件下, ω 越大,识别效率越低; ω 越小,识别效率越高。
召回率 ϱ	$\varrho = \frac{t_1}{t_1 + t_2}$

3.2.1 基于 Cifar-10 图像数据的目标识别实验

Cifar-10 数据集是由 Hinton 的学生 Alex Krizhevsky、Ilya Sutskever 收集的一个用于普通物体分类识别的计算机视觉数据集,它包含 60000 张 32×32 pixel 大小的 RGB 彩色图片,总共 10 个类别. 其中,包括 50000 张用于训练集,10000 张用于测试集,其图像部分样例如图 7 所示.

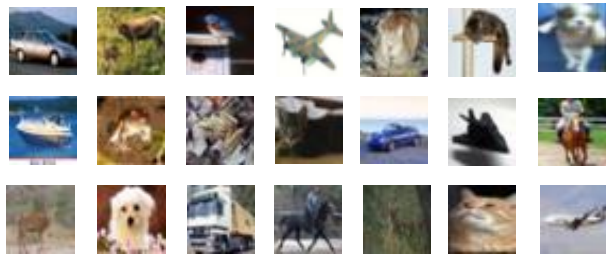


图 7 Cifar-10 数据集部分样例

Cifar-10 包含的是现实世界中真实的物体,但是噪声很大,而且物体的比例、特征都不尽相同,为实现分类识别带来很大困难,一般的线性模型在 Cifar-10 数据集上测试效果表现得很差. 因此,通过该数据集进行目标识别实验验证具备较高挑战性的同时,也可以更加充分的验证基于联合感知矩阵的压缩学习目标识别技术的可行性.

由于基于新型感知矩阵进行图像压缩的原始数据需为灰度图像,本实验将 Cifar-10 数据集也统一转化为灰度图作为实验数据. 实验选取了 Cifar-10 图像数据集中的飞机和汽车的各 5000 张数据进行图像识别实验验证,其中随机选取飞机和汽车共计 1000 张作为测试图像,其他图像作为训练集和验证集,验证集和训练集在模型训练过程中随机分配,其样式如图 8(a) 所示.

首先将 Cifar-10 数据集基于新型联合感知矩阵进行 4 倍压缩获取缩略图数据和观测值类图数据如图 8(a) 与 8(b) 所示.

将获取的缩略图数据和观测值类图数据进行拼接融合,可得到 Cifar-10 的多维压缩数据,其结果如图 9 所示.

基于图 6 构建的卷积神经网络进行实验验证,分别将 Cifar-10 原始图像数据集 κ_1 , 缩略图数据集 κ_2 , 以及构造的多维压缩数据集 κ_3 作为输入数据,在训练过程中学习率 $learning_rate = 0.0003$, 训练迭代次数

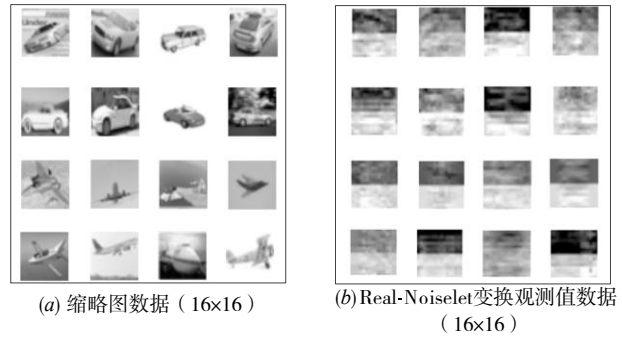


图 8 基于新型感知矩阵的 Cifar-10 数据变换结果图像

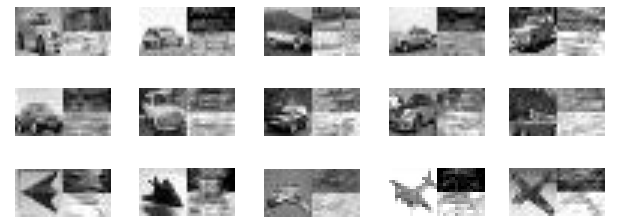


图 9 Cifar-10 多维压缩数据(24×12)

MAX_STEP = 10000, 训练获取的神经网络模型 Cifar-Model₁、Cifar-Model₂ 和 Cifar-Model₃ 通过对应的测试数据集进行测试,其中测试集图像总个数为 α_1 , 正确识别个数为 β_1 , 错误识别个数为 γ_1 , 训练时长为 δ_1 , 每 500 张目标图像从探测器生成并通过模型识别总时长 ω_1 . 通过对实验结果评价指标数据进行统计可以得到表 6.

通过对表 6 实验数据分析可知,本文提出的基于新型联合感知矩阵的压缩学习目标识别技术与基于原始图像数据集目标识别(Target recognition based on original image, TRO)方法相比,虽然识别精度略低,但是具备更高的识别与模型训练效率;与同等压缩比的基于缩略图数据集目标识别(Target recognition based on downsampled image, TRD)方法相比,目标识别与模型训练效率相近,但其识别准确率更高. 可得出基于联合感知矩阵的压缩学习目标识别技术与其他两种方法相比,在识别效果上不仅效率更高,并且与基于单空间域同等压缩比下的基于缩略图数据集目标识别方法相比,保持了更高的准确率.

3.2.2 基于 MNIST 手写图像数据集的目标识别实验

MNIST 数据集的来源是两个数据库的混合,一个

表 6 基于 Cifar-10 数据的目标识别实验数据统计表

指标 数据类型	训练时长 δ_1 (s)	测试图像个 数 α_1 (张)	正确识别 数目 β_1	错误识别 数目 γ_1	总时间 ω_1 (s)	识别准确率	飞机识别 召回率	汽车识别 召回率
κ_1	7364	1000	836	164	169.258	83.6%	90.9%	74%
κ_2	5268	1000	731	269	126.453	73.1%	78%	66.7%
κ_3	5474	1000	782	218	128.265	78.2%	78.5%	77.8%

来自 Census Bureau employees (SD-3), 一个来自 high-school students (SD-1). 该数据集有训练样本 60000 个, 测试样本 10000 个, 大小均为 28×28 pixel. 在训练样本和测试样本中, employee 和 student 写的都是各占一半. 60000 个训练样本一共由 250 个人书写, 其部分样例如图 10 所示.



图 10 MNIST 手写数据集部分样例

实验选取了 MNIST 数据集中 0 和 1 数据各 5000 张作为训练数据, 另选取测试集中 0 和 1 数据各 1000 张作为测试集数据, 并通过本文构造的目标识别神经网络进行目标识别效果验证.

首先进行多维压缩数据集的快速构建, 先将训练集与测试集通过新型联合感知矩阵进行 4 倍压缩获取缩略图和 Real-Noiselet 变换观测值数据图, 获取的缩略图和观测值数据部分样例如图 11(a) 与 11(b) 所示.

再基于缩略图数据与 Real-Noiselet 变换观测值数据图进行快速拼接融合, 生成多维压缩数据集, 其结果如图 12 所示.

通过上述构建的卷积神经网络分别对原始数据图像数据集 Ω_1 , 缩略图数据集 Ω_2 , 多维压缩数据图像数据集

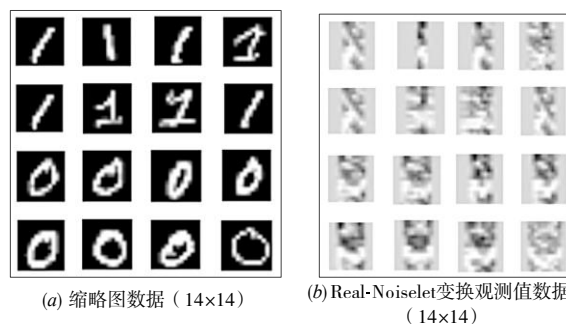


图 11 基于新型感知矩阵的 MNIST 数据变换结果图像

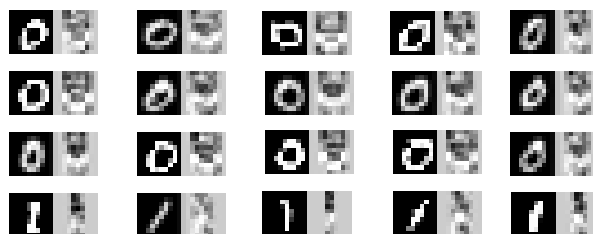


图 12 多维压缩数据图像(20 × 10)

Ω_3 三组数据进行实验验证. 在对不同实验数据验证过程中, 统一配置神经网络训练参数, 训练过程中学习率 $learning_rate = 0.0003$, 训练迭代次数 $MAX_STEP = 5000$. 并将训练获取的神经网络模型 Mini-Model₁、Mini-Model₂ 和 Mini-Model₃ 通过对应的测试集进行测试, 设其测试集图像总个数为 α_2 , 其中正确识别个数为 β_2 , 错误识别个数为 γ_2 , 训练时长为 δ_2 , 每 500 张目标图像从探测器生成并通过模型识别总时长为 ω_2 , 通过对实验结果评价指标数据进行统计可得到表 7.

表 7 基于 MNIST 数据的目标识别实验数据统计表

指标 数据类型	训练时长 δ_2 (s)	测试图像个 数 α_2 (张)	正确识别 目标个数 β_2	错误识别 目标个数 γ_2	总时间 ω_2 (s)	识别准确率	0 识别召回率	1 识别召回率
Ω_1	6984	1000	997	3	162.029	99.7%	99.8%	99.6%
Ω_2	4608	1000	986	14	123.568	98.6%	99%	98.2%
Ω_3	4817	1000	992	8	125.293	99.2%	99.4%	99%

通过对表 7 中实验结果数据分析可知, 基于新型联合感知矩阵的压缩学习目标识别技术在识别与模型训练效率方面均优于基于原始图像数据集目标识别方法, 与基于缩略图数据集目标识别方法几乎持平; 而且该方法的目标识别准确率也高于基于缩略图数据集目标识别方法, 虽然略低于基于原始图像数据集目标识别方法, 但仍然保持了较高的准确率. 综合可得出, 对比不同类型 MNIST 特征数据集进行的目标识别效果, 本文提出的基于新型联合感知矩阵的压缩学习目标识别技术不仅具备更高的识别效率, 而且保持了较高的识别准确率.

4 总结

本文提出的基于新型联合感知矩阵的压缩学习目标识别技术在传统空间域图像处理方式的基础上, 进行了扩展, 将图像的压缩域特征进行了充分的利用, 通过结合新型联合感知矩阵获取图像在压缩域和空间域的压缩数据, 进而生成新型多维压缩数据以代替原始图像数据实现对目标的高效高精度识别, 并以 Cifar-10 和 MNIST 数据集作为实验数据进行目标识别实验验证. 其验证结果表明, 该方法与传统深度学习目标识别方法相比具备更高的识别效率, 并且在同等压缩比下具备更高的识别准确率. 这种基于压缩码流拟直接实现目标识别的方式, 一方面提高了深度学习框架的数

据吞吐率,另一方面,可以降低数据的存储空间占用率,不仅提高了图像获取与压缩的效率,而且提升了目标识别效率,该方法特别适用于无需可视图像的场景,例如指纹识别与人脸识别等。

参考文献

- [1] Krizhevsky A, Sutskever I, GE Hinton. ImageNet classification with deep convolutional neural networks[J]. Communications of the ACM, 2017, 60(6):84 – 90.
- [2] Szegedy C, Liu W, Jia Y, et al. Going deeper with convolutions[A]. IEEE Conference on Computer Vision and Pattern Recognition[C]. Boston, Massachusetts: IEEE, 2015.1 – 9.
- [3] Lee S, Huh S, Yoo D, et al. Rich feature hierarchies from omni-directional RGB-DI information for pedestrian detection[A]. International Conference on Ubiquitous Robots and Ambient Intelligence[C]. Korea: IEEE, 2015. 362 – 367.
- [4] He K, Zhang X, Ren S, et al. Spatial pyramid pooling in deep convolutional networks for visual recognition[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2014, 37(9):1904 – 1916.
- [5] Ren S, He K, Girshick R, et al. Faster R-CNN: Towards real-time object detection with region proposal networks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39(6):1137 – 1149.
- [6] Redmon J, Divvala S K, Girshick R, et al. You only look once: Unified, real-time object detection[A]. IEEE Conference on Computer Vision and Pattern Recognition[C]. Palace: IEEE, 2016. 779 – 788.
- [7] Hang Du, Hailin Shi, Dan Zeng, Tao Mei. The elements of end-to-end deep face recognition: A survey of recent advances [A]. IEEE Conference on Computer Vision and Pattern Recognition[C]. Seattle, USA: IEEE, 2020.1 – 44.
- [8] Li H. Feature extraction, recognition, and matching of damaged fingerprint: Application of deep learning network[J]. Concurrency and Computation: Practice and Experience, 2021,33:E6057.1 – E6057.9.
- [9] Starck Jean-Luc, Candes E J, Donoho D L. The curvelet transform for image denoising[J]. IEEE Transactions on Image Processing, 2002, 11(6):670 – 684.
- [10] Said A, Pearlman W A. A new, fast, and efficient image codec based on set partitioning in hierarchical trees[J]. IEEE Transactions on Circuits and Systems for Video Technology, 1996, 6(3):243 – 250.
- [11] CANDES E J, ROMBERG J, TAO T. Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information[J]. IEEE Transactions on Information Theory, 2006,52(2):489 – 509.
- [12] Duarte M F, Davenport M A, et al. Single-pixel imaging via compressive sampling[J]. IEEE Signal Processing Magazine, 2008, 25:83 – 91.
- [13] Aes F M, Araújo F M, et al. Active illumination single-pixel camera based on compressive sensing[J]. Applied Optics, 2011, 50(4):405 – 414.
- [14] 焦李成, 杨淑媛, 刘芳,等. 压缩感知回顾与展望[J]. 电子学报, 2011, 39(7):1651 – 1662.
JIAO Li-cheng, YANG Shu-yuan, LIU Fang, et al. Development and prospect of compressive sensing[J]. Acta Electronica Sinica, 2011, 39(7):1651 – 1662. (in Chinese)
- [15] Tran D T, et al. Multilinear compressive learning[J]. IEEE Transactions on Neural Networks and Learning Systems, 2020, 32(4):1512 – 1524.
- [16] Davenport M A, Duarte M F, et al. The smashed filter for compressive classification and target recognition[A]. IEEE Conference on Computational Imaging V[C]. San Jose, CA: IEEE, 2007.64980H.1 – 64980H.12.
- [17] Lohit S, Kulkarni K, et al. Reconstruction-free inference on compressive measurements[A]. Conference on Computer Vision and Pattern Recognition Workshops [C]. Boston, Massachusetts: IEEE, 2015.16 – 24.
- [18] Jiangtao Wen, Zhuoyuan Chen, Yuxing Han, et al. A compressive sensing image compression algorithm using quantized DCT and noiselet information[A]. International Conference on Acoustics, Speech and Signal Processing [C]. Dallas, TX, USA: IEEE, 2010.1294 – 1297.
- [19] Tian Z, Luo C, Qiu J, et al. A distributed deep learning system for web attack detection on edge devices[J]. IEEE Transactions on Industrial Informatics, 2020, 16(3):1963 – 1971.
- [20] Lecun Y, Boser B E, Denker J S, et al. Backpropagation applied to handwritten zip code recognition[J]. Neural Computation, 1989, 1(4):541 – 551.
- [21] Glorot X, Bordes A, Bengio Y, et al. Deep sparse rectifier neural networks[A]. International Conference on Artificial Intelligence and Statistics[C]. USA: IEEE, 2011. 315 – 323.
- [22] Srivastava Nitish, Hinton Geoffrey, et al. Dropout: A simple way to prevent neural networks from overfitting[J]. Journal of Machine Learning Research, 2014, 15(56):1929 – 1958.
- [23] Rumelhart D, Hinton G, Williams R. Learning representa-

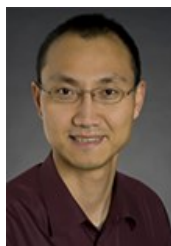
tions by back-propagating errors[J]. Nature, 1986, 323 (6088):533 – 536.

- [24] Liang X, Wang X, et al. Soft-margin softmax for deep recognition[A]. International Conference on Neural Information Processing[C]. Guangzhou, China: Springer, 2017. 413 – 421.

- [25] Krizhevsky A. Learning multiple layers of features from tiny images[D]. Department of Computer Science, University of Toronto, 2009.

- [26] Lecun Y, Bottou L, et al. Gradient-based learning applied to document recognition[J]. Proceedings of the IEEE, 1998, 86(11):2278 – 2324.

作者简介



李 峰 男,1975年生,吉林人,钱学森空间技术实验室研究员,主要研究方向为遥感图像处理与重建、压缩感知、探测效能评估等。
E-mail:lifeng@qxslab.cn



刘 洋 男,1992年生,山西忻州人,毕业于中国空间技术研究院西安分院,硕士研究生,主要研究方向为图像处理。
E-mail:liuandyang@aliyun.com



詹邦成(通信作者) 男,1995年生,河南信阳人.河南大学计算机与信息工程学院,硕士研究生,主要研究方向为目标识别、图像处理。
E-mail:zbc2862333729@dingtalk.com



刘志佳 男,1992年生,山西忻州人,毕业于中国空间技术研究院西安分院,硕士研究生,主要研究方向为图像处理。
E-mail:lzj303@sina.com



辛 蕾 女,1988年生,辽宁阜新,毕业于北京航空航天大学仪器科学与光电工程学院,硕士研究生,主要研究方向为压缩感知与空间信息智能提取。
E-mail:xinlei@qxslab.cn



肖化超 男,1985年生,湖南新邵人,毕业于中国空间技术研究院,工学博士,主要研究方向为空间数据处理。
E-mail:xiaohc@cast504.com