

基于神经网络与随机森林的涡旋光分类算法

张艳珠^{1,2} 黄浩铭^{1,2}

ZHANG Yanzhu HUANG Haoming

摘要

在光通信领域中,应用涡旋光可以大幅度提高光通信容量。但在实际的通信环境中,气候因素往往导致涡旋光经过散射介质形成散斑,大大影响涡旋光通信过程中的实际效果,而散斑图像识别分类涡旋光拓扑荷数对涡旋光作用在光通信中意义极大。将对两束涡旋光通过散射介质成像后的散斑图像进行分析,对涡旋光的拓扑荷数和相位信息进行分类。首先使用神经网络提取特征并进行机器学习,对两束涡旋光的拓扑荷数和相位信息等四个指标分别进行分类。然后尝试使用 ResNext 网络提取特征并向其中加入随机森林算法。最后实现了散斑图像中每个指标均超过 99% 的准确率以及同时满足四个指标 98.5% 的准确率。

关键词

图像分类; 涡旋光; 神经网络; 随机森林

doi: 10.3969/j.issn.1672-9528.2024.01.022

0 引言

近年来,神经网络在光学中辅助分类、去模糊、去湍流等领域应用频繁,在识别涡旋光拓扑荷数中也有所应用。随着人工智能^[1]的不断发展,机器学习被应用于拓扑荷数和对应相位^[2]的识别,测量涡旋光束的拓扑荷数和对应相位的光学转换和计算螺旋干涉条纹^[3]的问题已经被解决。

光纤在远距离传送^[4]信息方面有着广泛的应用,通过对其进行特殊的结构设计,可以保证其在传输过程中不发生畸变。涡旋光在光纤中传播时,会在光纤远端形成一幅杂乱无章的散斑图像,影响了拓扑荷数及对应相位的辨识。此外,本课题拟采用卷积神经网络与多目标分类器^[5]相结合的方法,在小样本的样本条件下,实现高精度的涡旋光束拓扑荷数与相位识别。提取基于 CNN 的特征^[6],将这些特征和相应的标记输入到多目标分类器中,最后完成分类。该方法在保证识别精度的前提下,可以大幅降低所需的数据量。

1 设计与方法

1.1 数据集采集与处理

对两束不同相位,不同拓扑荷数的涡旋光处理,共有两束涡旋光,初始状态都在相位为 1 的地方,且拓扑荷数也为 1,如图 1 中左图所示,两束涡旋光中拓扑荷数和相位信息更新和十进制更新变化相通变化顺序依次为第二束涡旋光拓扑荷

数、第二束涡旋光相位信息、第一束涡旋光拓扑荷数、第一束涡旋光相位信息。由此可见,两束涡旋光的相位信息和拓扑荷数都不相同,单个涡旋光相位信息为 10 种,拓扑荷数分类也为 10 类。两束涡旋光相互独立,共有 10^4 种情况,随机选择三张代表性的实际信息图如图 1 所示。

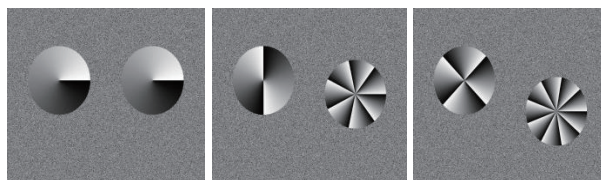


图 1 两束涡旋光实际信息图

两束涡旋光经过电路透过散射介质投射成散斑图像,不同拓扑荷数的涡旋光在不同相位投射后的散斑图像也随之不同,图 1 中三张图对应的散斑图像如图 2 所示。

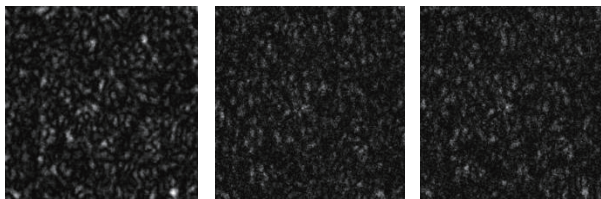


图 2 两束涡旋光散斑图像

1.2 网络架构

在获取实验数据后,为了保证在减少训练的数据样本的同时,提高准确率,本文尝试使用神经网络 ResNext101 及 ResNet101 对散斑图像提取深度特征,而后把提取到的特征送入随机森林进行拓扑荷数的分类及位置信息的分类,如图 3 所示。

1. 沈阳理工大学自动化与电气工程学院 辽宁沈阳 110000

2. 国家重点实验室控制学科研究生实验室 辽宁沈阳 110000

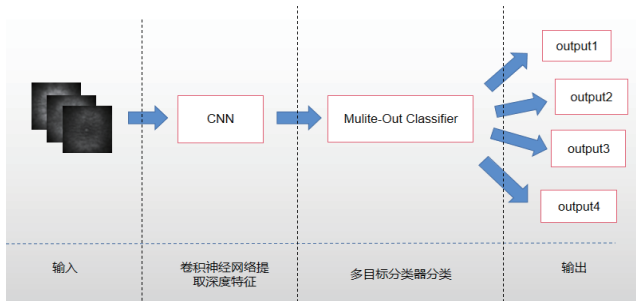


图3 基于神经网络和机器学习多输出分类器的网络架构

1.2.1 卷积神经网络

传统机器学习技术在提取图像特征的相关性和数量都比卷积神经网络提取图像特征更有局限性，因此，使用CNN神经网络对图像进行深度解析并提取特征，通过反向传递实现参数自动更新，从而相比于传统机器学习对图像特征提取更加准确，选择不同结构神经网络提取特征，包括GoogLeNet^[7]、ResNet^[8]和ResNext^[9]。

然而，训练一个深度比较深的深度神经网络非常难，在ResNet网络未开发之前的神经网络是不具备堆叠层数提高性能的。ResNet网络在Deep Learning领域中具有代表性，并且提出了残差结构，一个残差块有两条路径 $F(x)$ 和 x ， $F(x)$ 路径拟合残差，也称之为残差路径， x 路径为identity mapping恒等映射，称之为shortcut。图1中⊕为element-wise addition，要求参与运算的 $F(x)$ 和 x 尺寸相同。最初提及中，残差路径^[10]分为两种：一种有bottleneck结构，用于降维后再升维，主要为了便于降低计算复杂度的现实考虑；另一种则没有bottleneck结构。而shortcut路径也可以分为两种，取决于残差路径是否改变feature map的数量和尺寸如图4所示，左边残差结构用于ResNet-34，右边的残差结构主要用于ResNet-50/101等。

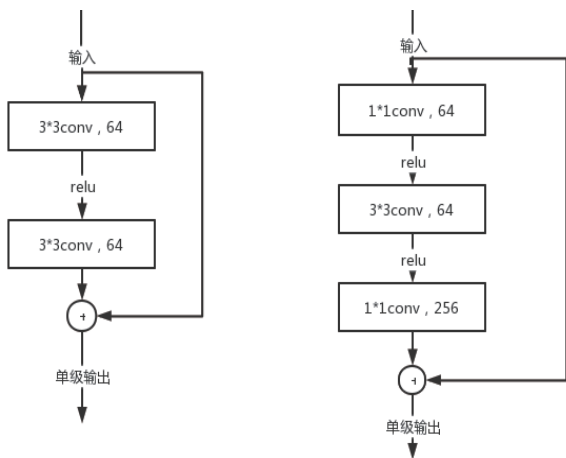


图4 两种主要的残差结构

ResNet-50/101网络是由用于降维或升维的残差块组成，如表1所示，列举了ResNet-34、50、101的网络结

构及参数，通过已知网络参数，可以建立相对效果较好的ResNet网络模型。

表1 ResNet网络结构

Layer name	Output size	ResNet-34 layer	ResNet-50 layer	ResNet-101 layer
Conv1	112×12	$7 \times 7, 64$ (stride 2)	$7 \times 7, 64$ (stride 2)	$7 \times 7, 64$ (stride 2)
Conv2_x	56×56	3×3 max pool (stride 2)	3×3 max pool (stride 2)	3×3 max pool (stride 2)
		$\begin{bmatrix} 3 \times 3, 64 \\ 3 \times 3, 64 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 64 \\ 3 \times 3, 64 \\ 1 \times 1, 256 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 64 \\ 3 \times 3, 64 \\ 1 \times 1, 256 \end{bmatrix} \times 3$
Conv3_x	28×28	$\begin{bmatrix} 3 \times 3, 128 \\ 3 \times 3, 128 \end{bmatrix} \times 4$	$\begin{bmatrix} 1 \times 1, 128 \\ 3 \times 3, 128 \\ 1 \times 1, 512 \end{bmatrix} \times 4$	$\begin{bmatrix} 1 \times 1, 128 \\ 3 \times 3, 128 \\ 1 \times 1, 512 \end{bmatrix} \times 4$
Conv4_x	14×14	$\begin{bmatrix} 3 \times 3, 256 \\ 3 \times 3, 256 \end{bmatrix} \times 6$	$\begin{bmatrix} 1 \times 1, 256 \\ 3 \times 3, 256 \\ 1 \times 1, 1024 \end{bmatrix} \times 6$	$\begin{bmatrix} 1 \times 1, 256 \\ 3 \times 3, 256 \\ 1 \times 1, 1024 \end{bmatrix} \times 23$
Conv5_x	7×7	$\begin{bmatrix} 3 \times 3, 512 \\ 3 \times 3, 512 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 512 \\ 3 \times 3, 512 \\ 1 \times 1, 2048 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 512 \\ 3 \times 3, 512 \\ 1 \times 1, 2048 \end{bmatrix} \times 3$
	1×1	Average pool, Fully connected layer(分类总数), softmax		
FLOPs (参数量需求)		3.6×10^9	3.8×10^9	7.6×10^9

从表1中信息可知，对于ResNet网络模型，默认输入图像尺寸为 224×224 ，即便不是，也可以通过图像预处理部分得到，首先都是由步长为2的 7×7 卷积层，然后通过核大小为 3×3 的最大池化层进行下采样操作，这样不仅可以减少计算量防止过拟合，还可以增加感受野，使后续进行的卷积操作可以学习到更全局的信息。ResNext网络相比于ResNet网络整体结构基本一致，区别在于ResNext将ResNet的残差模块分解为多个均匀的分支结构，这样设计可以使网络结构更加条理清晰，也可以同时减少对参数数量的需求，错误率越低。ResNext网络中残差结构如图5所示。

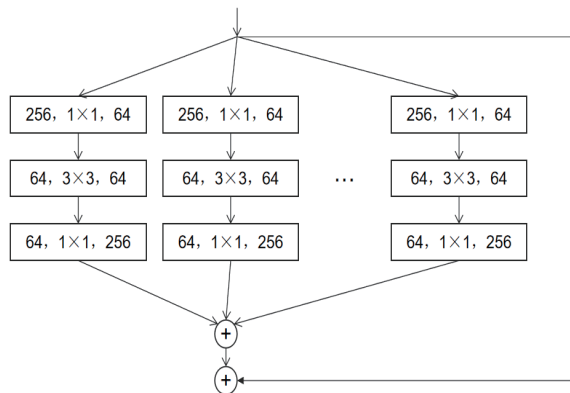


图5 ResNext网络残差结构

1.2.2 多目标分类器

本文网络架构通过 CNN 网络提取图像特征后送入架构中第二部分，多输出分类器。将在这部分实现对散斑图像的四个指标进行识别，并且在使用数据训练变少的情况下依然保留较高的识别准确率。由于在散斑图像中还原两束涡旋光拓扑荷数和其相位信息，实验具有四个目标值，因此需要对 CNN 提取出的特征送入多输出分类器^[11]，通过对评估器的类型选择达到对四个目标值的拟合、预测和分类。本次试验研究中第二部分使用的评估器是随机森林集成学习分类器，随机森林^[12]是一种将多棵树集成的一种集成学习方法，具有较高的准确率和抗干扰能力，并且支持高纬度的样本特征。结构如图 6 所示。

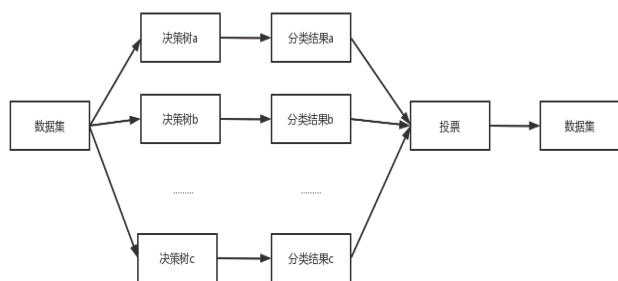


图 6 随机森林算法流程

2 实验结果

两束涡旋光的拓扑荷数^[13]从 1 到 10 交替变化，两束涡旋光的相位关系也从 1 到 10 交替变化，共有 10 000 张完全不一样的散斑图像。对 10 000 张图像随机选取作为网络的训练集、验证集和测试集，并对网络结构进行训练和测试。为了方便对散斑图像^[14]和其目标值对应，将每张图的 4 个目标值做出一一对应的 txt 文件作为标签送入网络，将带有标签的散斑图像送入网络。

本文首先尝试使用 8000 张随机散斑图像作为训练集，测试集为 2000 张散斑图像，DenseNet121 和 GoogLeNet 网络对四个指标的分类效果并不好，而 ResNet 网络对多数据量训练时可以学习并提取到有效特征，并且会随着网络深度，对参数的学习更精准。可以达到 95% 的总体准确率。但由于想对训练集数量进行适量缩减，重新对所有散斑图像进行随机分配，使用 4000 张随机散斑图像对网络结构进行训练，使用 6000 张图片进行测试，直接送入网络并测试指标准确率。使用 ResNet101 网络直接分类预测四个指标总体准确率最高，总体准确率仅仅达到 85.6%。

将缩减训练集后训练好的 CNN 网络稍加修改，使得网络作为提取图片深度特征的工具，再将特征经过处理为

张量，送入多目标分类器进行分类预测，结果如图 7 所示。由图 7 可知，使用 CNN-MOC 可以大幅度提升网络结构对四个指标的分类准确率，ResNext101-MOC 网络结构效果最佳，每个指标单独可以达到 99%。总体准确率可以达到 98.5%。

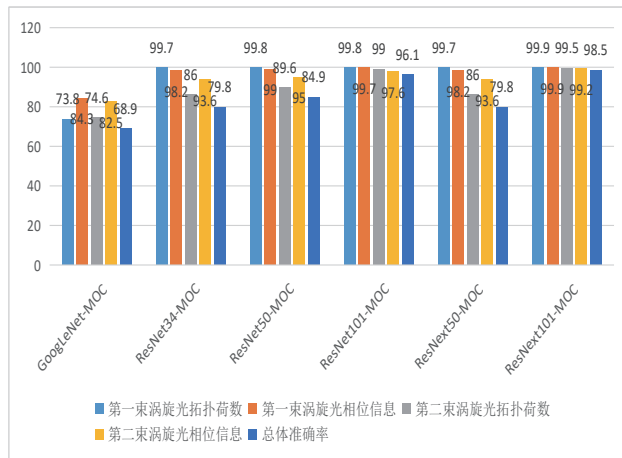


图 7 使用不同 CNN-MOC 效果对比图

3 结果与讨论

若将散斑图像训练集数量达到 8000 张，单独使用 CNN 网络对四个指标预测分类时，使用 ResNet101 网络可以相对其他网络更有效学习并准确分类，效果可以分别达到：99.8、99.8、97.8、97.5 和 95（第一束涡旋光拓扑荷数、第一束涡旋光相位信息、第二束涡旋光拓扑荷数、第二束涡旋光相位信息、总体准确率）。

想有效学习并提取有效特征，对图像的语义信息和分辨率要求很高，网络结构并不是增加深度就可以给效果带来提升，需要根据网络结构进行权衡分析。但对于 ResNet 和 ResNext 系列网络，增加网络深度可以学习到更深层次更有效的特征信息，通过对卷积模块的改变可以学习到更大分辨率中的特征信息。从而使用 ResNet101 网络总体准确率可以达到 95%，其他深度的 ResNet 和 ResNext 网络效果并不佳，都在 80% 左右，是因为 ResNet 网络模块结构可以识别到有效信息，但深度也可以提高网络性能和准确率。

但为了达到数据量少且实现效果佳的目的，直接使用 CNN 网络进行预测分类，当减少数据量使用 4000 张散斑图像训练，由于数据量的减少，学习到的特征有限，因此各个准确率大幅度降低，总准确率仅达到 85.6%。

由于减少网络训练数据量，从而导致学习到的有用信息特征减少，从而导致准确率降低。但将整体网络结构进行修改，将 CNN 神经网络作为提取深度特征的工具，将提取到

的特征信息转化为张量送入随机森林算法中,达到预测分类的效果。可以看到 CNN-MOC 的网络结构可以大幅度提升小样本数据集分类的准确率,增加网络深度和改进网络总体结构可以增强对图像在语义信息和分辨率两个方面对图像深度信息的学习和识别预测,总体效果得到很大程度的提升。四个指标的准确率均超过 99%,总体准确率达到 98.5%。因此本研究提出的基于深度学习和随机森林的涡旋光分类算法结构可以高精度、少数据达到高精度从两束涡旋光中分别提取拓扑荷数和相位信息。

4 结语

综上所述,提出一种将神经网络与多目标分类器算法结合的深度学习算法,将深度学习与传统机器学习算法相结合,使用 CNN 提取深度特征信息送入随机森林算法中实现分类预测。这样不仅降低深度学习对大量数据的依赖,也达到了传统机器学习对图像深度信息处理相对较差的缺陷。本文提出的方法既可以使用少量数据训练整体网络,也使得准确率达到 98.5%,为网络结构处理小数据样本问题提供一种新的方案。

参考文献:

- [1] 叶玉儿,李军依,曹萌,等.双模式涡旋光束的轨道角动量的精确识别[J].激光与光电子学进展,2021,58(18):1811021.
- [2] SOSKIN M S, GORSHLOV V N, VASNETSOV M V, et al. Topological charge and angular momentum of light beams carrying optical vortices[J]. Physical review A, 1997, 56(5): 4064.
- [3] GJR A A, HOFMANN A C, FANTIN A V, et al. An endoscopic optical system for inner cylindrical measurement using fringe projection[C]//Ninth International Symposium on Laser Metrology. New York:SPIE, 2008: 375-386.
- [4] POGGIOLINI P, POLETTI F. Opportunities and challenges for long-distance transmission in hollow-core fibres[J]. Journal of lightwave technology, 2022, 40(6): 1605-1616.
- [5] 邢昊,罗清,蔡和,等.偏振态对涡旋光束拓扑荷识别的影响[J].光学学报,2022,42(17):391-401.
- [6] LIN Z, JI K, KANG M, et al. Deep convolutional highway unit network for SAR target classification with limited labeled training data[J]. IEEE geoscience and remote sensing letters, 2017, 14(7): 1091-1095.
- [7] SZEGEDY C, LIU W, JIA Y, et al. Going deeper with convolutions[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2015: 1-9.
- [8] XIE S, GIRSHICK R, DOLLAR P, et al. Aggregated residual transformations for deep neural networks[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2017: 1492-1500.
- [9] FINZI M, BENTON G, WILSON A G. Residual pathway priors for soft equivariance constraints[J]. Advances in neural information processing systems, 2021, 34: 30037-30049.
- [10] SUEN C Y, LAM L. Multiple classifier combination methodologies for different output levels[C]//International Workshop on Multiple Classifier Systems. Berlin: Springer, 2000: 52-66.
- [11] BELGIU M, DRAGUT L. Random forest in remote sensing: A review of applications and future directions[J]. ISPRS journal of photogrammetry and remote sensing, 2016, 114: 24-31.
- [12] SOSKIN M S, GORSHKOV V N, VASNETSOV M V, et al. Topological charge and angular momentum of light beams carrying optical vortices[J]. Physical review A, 1997, 56(5): 4064.
- [13] CZERWINSHI R N, JONES D L, OBRIEN W D. Detection of lines and boundaries in speckle images-application to medical ultrasound[J]. IEEE transactions on medical imaging, 1999, 18(2): 126-136.
- [14] WaANG F, DEVABHAKTUNI V K, XI C, et al. Neural network structures and training algorithms for RF and microwave applications[J]. International Journal of RF and microwave computer - aided engineering, 1999, 9(3): 216-240.

【作者简介】

黄浩铭(1999—),男,辽宁鞍山人,研究生,研究方向:深度学习、机器学习结合对图像分类、图像去模糊等。

(收稿日期:2023-09-13)