# 基于坐标变换的四相涡旋光束轻量化识别方案

## 戈 峰,吴 凡,谢 超,李紫雯

上海理工大学光电信息与计算机工程学院,上海

收稿日期: 2025年4月30日; 录用日期: 2025年5月23日; 发布日期: 2025年5月30日

### 摘要

针对不同旋转角度的四相涡旋光束图像识别任务,文章提出了一种结合极坐标变换与Kolmogorov-Arnold 网络(KANs)的高效识别方法。传统卷积神经网络(Convolutional Neural Network, CNN)识别方法通常依赖 于高维数据的直接处理,不仅计算复杂度高,而且对旋转不变性特征的提取效果有限。而本文方法首先通 过极坐标变换,将笛卡尔坐标系下的二维光强分布图像映射为极坐标系下的一维特征序列。这种转换不仅 有效保留了光束的旋转对称性和角度特征,同时显著降低了数据维度,从而减少了后续计算的复杂性。 在此基础上,设计并构建了Kolmogorov-Arnold网络模型(KANs),充分利用其强大的特征提取和分类能 力,对经过极坐标变换的一维序列进行学习与分类。实验结果显示,该方法在面对不同旋转角度类别的 图像数据集中,表现出卓越的识别性能,识别准确率高达99.6%。同时,相较于传统方法,本方法显著 降低了时间复杂度和空间复杂度,提升了识别效率。本文提出的极坐标变换与KANs结合的识别框架为涡 旋光束的快速、高效识别提供了一种创新性解决方案,为光场信息的智能处理和应用拓展奠定了基础。

#### 关键词

四相涡旋光束,极坐标变换,Kolmogorov-Arnold网络,一维特征序列

# Four-Phase Vortex Beam Lightweight Identification Scheme Based on Coordinate Transformation

#### Feng Ge, Fan Wu, Chao Xie, Ziwen Li

School of Optical-Electrical and Computer Engineering, University of Shanghai for Science and Technology, Shanghai

Received: Apr. 30th, 2025; accepted: May 23rd, 2025; published: May 30th, 2025

#### Abstract

For the task of image recognition of four-phase vortex beams with different rotation angles, this paper

**文章引用:** 戈峰, 吴凡, 谢超, 李紫雯. 基于坐标变换的四相涡旋光束轻量化识别方案[J]. 建模与仿真, 2025, 14(5): 1079-1091. DOI: 10.12677/mos.2025.145458

proposes an efficient recognition method combining polar coordinate transformation and Kolmogorov-Arnold networks (KANs). The traditional convolutional neural network (CNN) recognition method usually relies on the direct processing of high-dimensional data, which not only has high computational complexity, but also has a limited effect on the extraction of rotational invariant features. In contrast, the method in this paper first maps the two-dimensional light intensity distribution image in a Cartesian coordinate system to a one-dimensional feature sequence in a polar coordinate system through polar coordinate transformation. This transformation not only effectively preserves the rotational symmetry and angular features of the light beams, but also significantly reduces the data dimensions, which in turn reduces the complexity of the subsequent calculations. On this basis, Kolmogorov-Arnold network models (KANs) are designed and constructed to make full use of their powerful feature extraction and classification capabilities to learn and classify the one-dimensional sequences transformed by polar coordinates. Experimental results show that the method exhibits excellent recognition performance in the face of image datasets with different rotation angle categories, with recognition accuracy as high as 99.6%. Meanwhile, compared with the traditional method, this method significantly reduces the time complexity and space complexity and improves the recognition efficiency. The recognition framework combining polar coordinate transforms and KANs proposed in this paper provides an innovative solution for the fast and efficient recognition of vortex beams and lays the foundation for the intelligent processing and application expansion of light field information.

## **Keywords**

Four-Phase Vortex Beam, Polar Coordinate Transformation, Kolmogorov-Arnold Network, One-Dimensional Feature Sequence

Copyright © 2025 by author(s) and Hans Publishers Inc. This work is licensed under the Creative Commons Attribution International License (CC BY 4.0). <u>http://creativecommons.org/licenses/by/4.0/</u>

CC Open Access

# 1. 引言

涡旋光束是一种携带轨道角动量(Orbital Angular Momentum, OAM)的特殊光束,其波前呈螺旋状分 布,具有独特的相位奇点和光强分布特性[1]。由于 OAM 模式在理论上具有无限维度,涡旋光束在高容 量光通信[2]、量子信息处理[3]、光学微操纵[4]等领域展现出巨大的应用潜力。然而,涡旋光束的拓扑荷 数(即 OAM 模式)识别是实现其应用的关键技术之一。传统的识别方法主要依赖于干涉仪[5]、衍射光栅[6] 或计算全息图[7]等装置,这些方法虽然能够实现较高的识别精度,但往往需要复杂的实验装置和较高的 计算成本,难以满足实时性和便携性需求。对于不同旋转角度的四相涡旋光束,传统方法在处理其二维 光强分布时面临双重挑战:一方面,二维图像的高维度特性导致计算复杂度呈指数级增长;另一方面, 光束旋转角度变化会引入额外的空间畸变,显著降低识别鲁棒性[8][9]。

近年来,随着深度学习技术的快速发展,基于神经网络的光学模式识别方法逐渐成为研究热点[10]。 与传统方法相比,神经网络能够通过端到端的学习方式自动提取特征,显著降低了人工设计特征的复杂 性[11]。然而,现有的基于卷积神经网络(CNN)的方法在处理高维光学数据时,仍然面临两个关键瓶颈: 其一,二维卷积核在提取旋转敏感特征时存在局限性,导致模型对光束角度的变化适应性不足[12];其二, 处理涡旋光强图像需要多层卷积堆叠,显著增加了时间复杂度和硬件资源消耗[13]。这些局限性在需要实 时处理多角度四相涡旋光束的应用场景中尤为突出。

本文提出了一种基于极坐标变换的数据降维方法,显著降低了涡旋光束识别的计算复杂度,并增强

了对不同旋转角度四相涡旋光束的鲁棒性;其次,构建了 KANs 网络模型,实现了对极坐标变换后一维 序列的高效准确识别;最后,通过实验验证了该方法在识别准确率、时间复杂度和空间复杂度方面的优 越性。实验结果表明,该方法在识别准确率达到 99.6%的同时,显著降低了计算资源消耗,为涡旋光束的 快速识别提供了一种新思路。

## 2. CNN 网络的结构

CNN 的神经元之间是局部连接的,并组织成层级架构,在相同的参数下,经过前一层的网络不同位置进行计算,从而得到一种平移不变性的网络结构,通常用于处理图像、视频和声音等数据。CNN 基本结构包括卷积层(Convolutional Layer)、池化层(Pooling Layer)、全连接层(Fully Connected Layer)和激活函数(Activation Function)等[14] [15]。

#### 2.1. 卷积层

卷积层是卷积神经网络中最关键的组成部分,其主要功能是从输入数据中提取有意义的特征。卷积 层的核心由多个卷积核组成,每个卷积核通过与输入数据局部区域进行逐元素计算,提取出特定的特征 信息。具体而言,卷积核在输入数据上以滑动窗口的方式逐点移动,计算局部区域的加权和,并将计算 结果作为输出特征图的一部分。最终,卷积核完成对整个输入数据的扫描后,生成完整的输出特征图。

在卷积层中,神经元被组织为特征图的形式,每个神经元只与输入数据的局部区域相连接,这种连 接方式被称为局部连接。这种设计基于图像数据的统计特性,图像中的相邻像素通常具有更高的相关性, 而远距离像素的相关性较低。因此,局部连接能够有效专注于提取局部特征,如边缘、纹理和角点等, 同时显著减少计算复杂度和参数数量。

卷积运算将输入数据的局部区域信息编码为输出特征图的元素值,从而提取到关键的模式或特征。 例如,对于一个大小为2×2的卷积核在输入图像上滑动时,每次操作会对局部的2×2区域进行加权求 和,最终得到一个3×3的特征图。如图1所示,展示了卷积计算的具体过程。





## 2.2. 池化层

池化层是卷积神经网络中的另一重要组成部分,通常位于两个卷积层之间,其主要功能是对特征图 的尺寸进行压缩,同时保留关键特征信息,为后续层的计算提供更紧凑的表示。通过池化操作,可以有 效地减少参数数量和计算量,降低过拟合风险,同时增强模型对输入数据平移的鲁棒性。池化操作的核 心思想是通过对输入特征图的局部区域进行聚合,生成尺寸缩小的输出特征图,两种最常见的池化方式 为最大池化(Max Pooling)和平均池化(Average Pooling)。最大池化下采样层,在输入数据的每个小区域内 取最大值作为该区域的池化输出,能够保留输入数据的重要特征。平均池化下采样层,在输入数据的每 个小区域内取平均值作为该区域的池化输出,生成平滑的特征图。

假设在一个4×4的特征图上使用2×2的池化窗口进行最大池化操作,每次滑动步长为2,则输出 特征图的尺寸将变为2×2,这种降维能够显著减少数据量,同时保留输入特征图的核心信息。图2展示 了最大池化计算过程。



**Figure 2.** Algorithms for maximum pooling downsampling and average pooling downsampling 图 2. 最大池化下采样和平均池化下采样的运算过程

#### 2.3. 全连接层和激活函数

全连接层是卷积神经网络中关键的决策模块,通常用于对卷积和池化层提取到的特征进行分类或回 归运算。其基本思想是将前一层的特征图展平成一维向量,并通过权重矩阵计算各类别或回归目标的预 测值。由于全连接层与传统的神经网络结构一致,其每个神经元与上一层的所有神经元相连,因此全连 接层的参数量较大。

为了使神经网络能够拟合复杂的非线性关系,通常在卷积层或全连接层的后面使用激活函数,以此 对线性变换的结果进行非线性映射,从而增强网络的表达能力。常用的激活函数有线性整流函数 ReLU, 当 ReLU 函数的输入 *x* 小于等于 0 时,其输出为 0,因此可以将一些不重要的特征直接置为 0,从而增强 模型的稀疏性,减少过拟合的风险。ReLU 激活函数的表达式为:

$$f(x) = \operatorname{Max}(0, x) \tag{1}$$

在针对多分类任务处理时,常用的激活函数为 Softmax,用于将神经网络的输出值转换为概率分布,并匹配最大概率的类别为输出的预测结果,处理后的所有输出节点概率之和为 1,通常与交叉熵损失函数联合使用。Softmax 函数的表达式为:

$$Softmax(y_i) = \frac{e^{x_i}}{\sum_i^N e^{x_i}}$$
(2)

#### 2.4. ResNet-50 网络

本文选用 ResNet-50 作为传统 CNN 识别的模型, 其具有 50 层深度, 包含 49 层卷积层和 1 层全连接

层,其架构是由多个残差模块组成[16]。共分为5个阶段(Stage 0 至 Stage 4)。输入(Input)经过 ResNet-50 的5个阶段得到输出(Output), Stage 0 结构是对输入的预处理,后面的 Stage 都是由若干个瓶颈结构(Bottle-neck)组成。

Bottleneck 是由 3 层卷积操作完成,此过程可以大大减少计算量,提高运算效率。首先 1 × 1 的卷积 层是减少通道数量来压缩特征维度;再通过 3 × 3 卷积层负责提取特征,由于图像已经在前一层被降维操 作,计算量已经明显降低;最后一层 1 × 1 的卷积层用于扩展通道数,恢复特征维度,使得输入和输出维 度相同。

在这些残差模块中,通过引入直连或跳跃连接将输入与权重层的输出相加,使得网络可以学习到残 差函数,从而避免梯度消失问题,有助于网络训练,其结构如图3所示。



Figure 3. ResNet structural framework 图 3. ResNet 结构框架

## 3. KANs 网络架构与仿真模型

通过在高斯光束中引入偏振片与涡旋半波片对光场的偏振态与相位进行联合调制[17],可实现具有 四个离散相位奇点的拉盖尔 - 高斯(Laguerre Gaussian, LG)模式光束,即四相涡旋光束[18]。其相位分布可 解析为以下两部分:一是携带轨道角动量(OAM)的螺旋相位的涡旋因子。二是将光束横截面分为四个象 限的四相分区调制,每个象限赋予不同的相位或拓扑荷数。总相位的表达式为:

$$\phi(x, y) = \sum_{n=1}^{4} P_n \cdot R_n(x, y)$$
(3)

其中, R<sub>n</sub>(x, y)是第 n 个象限的矩形函数, P<sub>n</sub>是各象限的相位调制值(0, π/2, π, 3π/2 或其它值)。

引入高斯噪声模拟实验噪声,如探测器热噪声、环境扰动等。噪声标准差  $\sigma = I_{\text{max}}/SNR$ , 信噪比 SNR = 10 dB。

$$I_{noisy} = I_{ideal} + N(0, \sigma^2)$$
(4)

Kolmogorov-Arnold 网络[19],也被称为 KANs,源自 Kolmogorov 与 Arnold 关于函数分解的数学定

理。该定理指出,任意多元连续函数 f(x)均可表示为有限个单变量函数,连续函数的两层嵌套加法的形式,即任何多变量连续函数都可以表示为单变量连续函数和加法运算的组合,使得神经网络能够有效地学习和训练,其数学形式为:

$$f(x_1, x_2, ..., x_n) = \sum_{j=1}^{2n+1} g_j\left(\sum_{i=1}^n \phi_i(x_i)\right)$$
(5)

其中, $\phi_i(x_i)$ 是输入变量的非线性变换, $g_j$ 是线性或非线性加权函数,2n+1是所需的函数组合数量,具体数量与输入维度n相关。

针对四相涡旋光束一维序列的复杂模式识别需求,本文基于高效 KANs 网络框架,通过系统性实验 探索网络深度对分类性能的影响。实验采用网格搜索策略,在保持总参数量(约 72.5 M)相近的条件下,对 2 至 10 层网络架构进行对比分析,如表 1 所示。

# Table 1. Data results for different network architectures 表 1. 不同网络架构数据结果

Layers	Accuracy/%	Training time/s	Configuration
2	96.6	71.03	[512, 512, 10]
3	97.6	83.06	[512, 512, 512, 10]
4	98.9	100.07	[512, 512, 512, 512, 10]
5	99.6	120.01	[512, 512, 512, 512, 512, 10]
6	99.5	139.10	$[512^6, 10]$
8	99.0	173.03	$[512^8, 10]$
10	98.1	215.05	[512 <sup>10</sup> , 10]

在2至5层区间内,准确率随深度增加提升(+3.0%)。当层数在5层时,精度接近饱和。而超过5层 后,模型出现轻微过拟合现象,验证损失上升 0.15,因此证明四相涡旋光特征识别需要足够的非线性映 射能力。选择5层架构作为最优配置,其准确率和效率平衡点满足实时识别需求。

## 4. 图像采集

对四相涡旋光束进行采集,光路图如图4所示。

本实验的设计旨在为涡旋光束的生成与高效识别提供稳定的实验基础,通过实验系统采集其光强分 布图像,为后续的涡旋光束特征提取与识别提供数据支撑。系统组成包括以下关键部分:稳定的激光光 源、用于准直和聚焦光束的光学凸透镜、控制光束偏振态的线偏振片、将高斯光束转换为具有四个相位 奇点的涡旋光束的涡旋半波片,以及用于控制实验设备和存储采集图像的计算机。

实验中,激光器发射波长为 633 nm 的线偏振高斯光束,作为光源的初始输入。通过平凸透镜(焦距 22 cm、直径 25.4 mm)对高斯光束进行准直和发散处理,以增加光束的覆盖范围。随后,光束进入第一块 线偏振片(间距 7 cm、直径 25.4 mm,偏振角度固定为 0°),通过控制光束的偏振态生成高质量的线偏振 高斯光束,作为涡旋光束生成的基础。接下来,光束穿过涡旋半波片(间距 5 cm、直径 25.4 mm),该光学 元件通过对光束的偏振态和相位进行联合调制,将输入的高斯光束转换为具有四个相位奇点的拉盖尔 - 高斯(Laguerre-Gaussian, LG)光束。为进一步优化光束特性,在实验系统中加上了第二块线偏振片(间距 7 cm、直径 25.4 mm,偏振角度可调),对光束的偏振态进行二次调制,从而去除多余的偏振分量,确保涡

旋光束的纯度和成像质量,偏振片的角度以10°为步长递增旋转,共完成72次光束偏振调制。



Figure 4. Image acquisition optical path diagram 图 4. 图像采集光路图



**Figure 5.** Optical whiteboard imaging 图 5. 光学白板成像情况



Figure 6. Four-phase vortex beam image 图 6. 四相涡旋光束图像

调制后的涡旋光束射向距离偏振片 35 cm 的光接收板,利用接收板记录光强分布图像,并根据光束成 像质量,适当调整光源位置和光学元件的相对位置,图 5 为光学白板成像情况。将采集的光强分布图像存 储到计算机中,作为后续分类与识别的输入数据,并部署深度学习模型,对采集到的数据进行分类与特征 提取。经过实验生成不同旋转角度的四相涡旋光束图像,手动裁剪初步生成图像,把图像以外的冗余位置 信息去除掉,如图 6 所示。可以清晰观察到光束的四个相位奇点,其分布呈现出明显的旋转对称特性。

## 5. 极坐标变换与图像信息压缩

本研究的核心是通过极坐标变换将笛卡尔坐标系下的二维图像 *I* (*x*, *y*)转换为极坐标系下的 *I* (*r*, *θ*), 实现二维光强分布到一维特征序列的压缩[20],同时保留光场旋转对称性和角度特征。针对像素为 256× 256 的四相涡旋光束灰度图像,具体变换方式如下:

$$r = \sqrt{x^2 + y^2} \tag{6}$$

$$\theta = \tan^{-1} \left( \frac{y}{x} \right) \tag{7}$$

其中 r 表示像素点(x, y)在笛卡尔坐标系中到质心(0, 0)的距离, θ 表示像素点与质心形成的向量(x, y)与向量(1,0)之间的夹角。经过坐标变换后,原始图像若得到质心坐标为(X<sub>0</sub>, Y<sub>0</sub>),通过公式(4)、(5)得出极坐标的转换公式如下:

$$r = \sqrt{\left(x - X_0\right)^2 + \left(y - Y_0\right)^2}$$
(8)

$$\theta = \tan^{-1} \left( \frac{y - Y_0}{x - X_0} \right) \tag{9}$$

通过以上公式成功完成图像的极坐标变换处理,如图7所示。



**Figure 7.** The original image is grayed out and then transformed by polar coordinates 图 7. 原图像经过灰度处理后,再通过极坐标变换后的图像

极坐标变换过程中,通过双线插值法来确定图像变换后的插值点[21],如图 8 所示。它通过利用周围 像素的值来估计图像中任意位置的像素值。通过四个最近邻近像素值来计算插值点的值,假设需要插值 的位置为(*x*, *y*),其周围四个邻近像素为(*x*<sub>1</sub>, *y*<sub>1</sub>)、(*x*<sub>1</sub>, *y*<sub>2</sub>)、(*x*<sub>2</sub>, *y*<sub>1</sub>)和(*x*<sub>2</sub>, *y*<sub>2</sub>),对应的像素值分别为*I*<sub>11</sub>、*I*<sub>12</sub>、 *I*<sub>21</sub>和*I*<sub>22</sub>。插值计算公式如下:

在水平方向上对目标上下两行进行线性插值:

$$R_{1} = \frac{(x_{2} - x)}{(x_{2} - x_{1})} I_{11} + \frac{(x - x_{1})}{(x_{2} - x_{1})} I_{21}$$
(10)

戈峰

$$R_{2} = \frac{(x_{2} - x)}{(x_{2} - x_{1})} I_{12} + \frac{(x - x_{1})}{(x_{2} - x_{1})} I_{22}$$
(11)

在垂直方向上对插值 R1、R2进行线性插值,得到最终像素值:

$$P_{(x,y)} = \frac{(y_2 - y)}{(y_2 - y_1)} R_1 + \frac{(y - y_1)}{(y_2 - y_1)} R_2$$
(12)

坐标变换处理后,将图像在水平方向和垂直方向上的像素灰度值进行平均。随后计算了图像每列的 平均灰度值,得到了 256 个值的一维序列,为了减少图像右侧不规则阴影处对识别精度的影响,在计算 图像的行平均值之前切断了右侧 50 列,因此计算出另一个 256 值的一维序列。最后,将两个序列拼接在 一起,得到了 512 个值的一维序列。





### 6. 结果分析

如表 2 所示,本文的程序是在配备 16 GB 内存的 NVIDIA Geforce RTX-3060 GPU 平台上训练的,所 使用的实验机器操作系统是 Window 10,所提出的算法是在 PyTorch 1.7.0 上构建的。

实验构建了两个数据集,分别为覆盖旋转角度  $\theta \in [0^\circ, 540^\circ]$ 的原始四相涡旋光束图像数据集和经过 坐标变换处理后的图像数据集。旋转角度按照每 60°间隔划分为 10 个子组,每组包含 800 张训练图像、 100 张验证图像和 100 张测试图像,共计 10,000 张样本。为了模拟实际信号环境,数据集在信号传输链 路中注入高斯白噪声,信噪比设置为 *SNR* = 10 *dB*。

针对原始图像数据集,训练周期设置为60个epoch,批处理尺寸(batch size)固定为32,学习率为1e-4。

对于经过坐标变换处理后的图像样本,将其转换为 512 维的一维序列表示,训练周期设置为 30 个 epoch, 批处理尺寸同样固定为 32, 学习率为 1e-4。

实验结果表明,轻量化识别方法实现了 99.6%的识别准确率,如图 9 所示。KANs 在处理极坐标变换 后的一维序列任务时,有着可观的精准度,且收敛速度极快。





**Figure 9.** Accuracy in samples after training polar transformations for KANs 图 9. KANs 训练极坐标变换后样本中的准确率

在传统识别方法中,用 ResNet-50 网络直接训练四相涡旋光束图像数据集,结果准确率可以达到 99.5%, 如图 10 所示。但是传统方法消耗的时间和资源成本较大,本文方法可将传统训练时间从 75 分钟压缩至 2 分钟,峰值显存占用从 8.6 GB 降低至 1.6 GB,表 3 给出了关键指标对比。传统识别与本次轻量化识别方法相比,虽然准确率相当,但在其它关键指标对比下,如时间复杂度和空间复杂度上有着显著差距。

这种改进确保了在处理四相涡旋光束图像的工作中,保持准确率的同时显著提升工作效率,也对硬件条件有了更多选择。通过对测试集的混淆矩阵进行错误分析,错误样本均出现在相邻角度数据组。分析原因发现,由于旋转角度间隔较小,导致图像特征性相似。所有错误都发生在相邻组内,未出现跨组误判,表明了模型对图像  $\Delta\theta = 10^{\circ}$ 的角度有着较强的区分能力。测试集的混淆矩阵如图 11 所示。



Figure 10. Accuracy of ResNet50 training a four-phase vortex beam image dataset 图 10. ResNet50 训练四相涡旋光束图像数据集的准确率

Table	e <b>3.</b> Comparison	of key metric	s between the	e methods in	this paper a	and traditional	training	methods
表 3.	本文方法与传	统训练方法主	:键指标对比	,				

Targets	Traditional Training	This Approach	Upgrade Ratio		
Accuracy	99.5%	99.6%	+ 0.1%		
Training time	75 min	2 min	↓ 97.3%		
Memory	8.6 GB	1.6 GB	↓ 81.4%		

Predicted Angle													
		°	ĥ	200	,0°	240	ŝ	Å,	82°	<sup>b</sup> <sup>®</sup> °	54°		
	0°	- 100	0	0	0	0	0	0	0	0	0	10	
	50°	- 0	99	1	0	0	0	0	0	0	0	10	U
12	20°	- 0	0	100	0	0	0	0	0	0	0	- 80	)
18	30°	- 0	0	0	99	1	0	0	0	0	0		
el 240 pur el 240 300 360 420 480	40°	- 0	0	0	0	100	0	0	0	0	0	- 60	,
	0°	- 0	0	0	0	0	100	0	0	0	0	- 40	į
	50°	- 0	0	0	0	0	0	100	0	0	0		
	20°	- 0	0	0	0	0	0	1	99	0	0	- 20	)
	30°	- 0	0	0	0	0	0	0	0	99	1		
54	10°	- 0	0	0	0	0	0	0	0	0	100		

Figure 11. Confusion matrix for the test set 图 11. 测试集的混淆矩阵

## 7. 结论

本文提出了一种基于 Kolmogorov-Arnold 网络(KANs)的四相涡旋光束一维序列识别方法,通过极坐标变换将二维光强分布映射为一维序列,显著降低了数据维度,并结合 KANs 的轻量化网络架构,在保证 99.6% 高识别精度的同时,将传统 CNN 网络训练识别的时间从 75 分钟压缩至 2 分钟(效率提升 97.3%),显存占用从 8.6 GB 降至 1.6 GB (空间复杂度降低 81.4%)。混淆矩阵分析表明,错误样本集中在相邻角度 ( $\Delta \theta = 10^\circ$ ),跨组误判率为零,验证了模型对角度差异的严格区分能力。尽管本研究取得了显著进展,但仍存在一些挑战,当前方法对  $\Delta \theta \leq 5^\circ$ 的图像区分效果还有提升空间,且在强噪声环境下,测试集准确率下降 5%以上,需进一步引入抗噪声模块。

# 参考文献

- Allen, L., Beijersbergen, M.W., Spreeuw, R.J.C. and Woerdman, J.P. (1992) Orbital Angular Momentum of Light and the Transformation of Laguerre-Gaussian Laser Modes. *Physical Review A*, 45, 8185-8189. https://doi.org/10.1103/physreva.45.8185
- [2] Willner, A.E., Huang, H., Yan, Y., Ren, Y., Ahmed, N., Xie, G., et al. (2015) Optical Communications Using Orbital Angular Momentum Beams. Advances in Optics and Photonics, 7, 66-106. <u>https://doi.org/10.1364/aop.7.000066</u>
- [3] Erhard, M., Fickler, R., Krenn, M. and Zeilinger, A. (2017) Twisted Photons: New Quantum Perspectives in High Dimensions. *Light: Science & Applications*, 7, 17146-17146. <u>https://doi.org/10.1038/lsa.2017.146</u>
- [4] Grier, D.G. (2003) A Revolution in Optical Manipulation. *Nature*, **424**, 810-816. <u>https://doi.org/10.1038/nature01935</u>
- [5] Berkhout, G.C.G., Lavery, M.P.J., Courtial, J., Beijersbergen, M.W. and Padgett, M.J. (2010) Efficient Sorting of Orbital Angular Momentum States of Light. *Physical Review Letters*, **105**, Article ID: 153601. <u>https://doi.org/10.1103/physrevlett.105.153601</u>
- [6] Lavery, M.P.J., Speirits, F.C., Barnett, S.M. and Padgett, M.J. (2013) Detection of a Spinning Object Using Light's Orbital Angular Momentum. *Science*, 341, 537-540. <u>https://doi.org/10.1126/science.1239936</u>
- [7] Xie, G.D., Ren, Y.X., Huang, H., et al. (2013) Analysis of Aperture Size for Partially Receiving and De-Multiplexing 100-Gbit/s Optical Orbital Angular Momentum Channels over Free-Space Link. 2013 IEEE Globecom Workshops (GC Wkshps), Atlanta, 9-13 December 2013, 1116-1120. <u>https://doi.org/10.1109/glocomw.2013.6825142</u>
- [8] Forbes, A., Dudley, A. and McLaren, M. (2016) Creation and Detection of Optical Modes with Spatial Light Modulators. *Advances in Optics and Photonics*, **8**, 200-227. <u>https://doi.org/10.1364/aop.8.000200</u>
- [9] Wang, J., Yang, J., Fazal, I.M., Ahmed, N., Yan, Y., Huang, H., et al. (2012) Terabit Free-Space Data Transmission Employing Orbital Angular Momentum Multiplexing. *Nature Photonics*, 6, 488-496. <u>https://doi.org/10.1038/nphoton.2012.138</u>
- [10] LeCun, Y., Bengio, Y. and Hinton, G. (2015) Deep Learning. *Nature*, **521**, 436-444. <u>https://doi.org/10.1038/nature14539</u>
- [11] He, K., Zhang, X., Ren, S. and Sun, J. (2016) Deep Residual Learning for Image Recognition. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, 27-30 June 2016, 770-778. https://doi.org/10.1109/cvpr.2016.90
- [12] Krizhevsky, A., Sutskever, I. and Hinton, G.E. (2012) Imagenet Classification with Deep Convolutional Neural Networks. *Communications of the ACM*, 60, 84-90.
- [13] Simonyan, K. and Zisserman, A. (2014) Very Deep Convolutional Networks for Large-Scale Image Recognition.
- [14] 周飞燕,金林鹏,董军.卷积神经网络研究综述[J]. 计算机学报, 2017, 40(6): 1229-1251.
- [15] Goodfellow, I., Bengio, Y. and Courville, A. (2016) Deep Learning. MIT Press.
- [16] Yang, Z., Xiong, F., Pei, Y., Chen, Z., Zhan, C., Hu, E., *et al.* (2024) A Stay Cable Icing Identification Method Based on the Fusion of U-Net and Resnet50. *Applied Sciences*, 14, 3062-3077. <u>https://doi.org/10.3390/app14073062</u>
- [17] Cardano, F. and Marrucci, L. (2015) Spin-Orbit Photonics. *Nature Photonics*, 9, 776-778. <u>https://doi.org/10.1038/nphoton.2015.232</u>
- [18] Zhan, Q. (2009) Cylindrical Vector Beams: From Mathematical Concepts to Applications. Advances in Optics and Photonics, 1, 1-57. <u>https://doi.org/10.1364/aop.1.000001</u>

- [19] Liu, Z., Wang, Y., Vaidya, S., et al. (2024) Kan: Kolmogorov-Arnold Networks.
- [20] Tsai, Y.C. and Chou, T.N. (2018) Deep Learning Applications in Pattern Analysis Based on Polar Graph and Contour Mapping. *International Journal of Applied Science and Engineering*, **15**, 183-198.
- [21] Gonzalez, R.C. (2009) Digital Image Processing. Pearson Education India.