

基于张量低管道秩的图像多分类模型

张家瑞, 胡毓榆, 唐开煜, 樊亚莉*

上海理工大学理学院, 上海

收稿日期: 2024年4月29日; 录用日期: 2024年5月23日; 发布日期: 2024年5月31日

摘要

传统机器学习方法在对高阶张量数据进行分类时, 往往将其转化为低阶格式, 由此会产生过拟合问题并且破坏张量的结构。针对上述问题提出一种基于张量低管道秩的多分类模型(LRTMLR)。该模型可以直接对张量格式的图像进行分类, 使用由张量-张量积诱导的张量管道秩及相应的张量核范数来处理低秩张量, 更好地利用张量结构特点, 提高张量格式图像的多分类准确性。在三分类仿真数据集上, LRTMLR的分类准确率较无结构信息(MLR)、带矩阵结构信息(LRMLR)的方法均提升9.6个百分点, 在五分类仿真数据集上则分别提升23.2和25.2个百分点。在加州理工大学的101类彩色图像识别数据集的三分类、五分类和十四分类子集上, LRTMLR的分类准确率较MLR分别提升了10.01、25.61和40.78个百分点, 较LRMLR分别提升了10.68、25.61和40.78个百分点, 与基于CP分解的方法(MCPLR)相比提高了6.47、13.37和27.73个百分点, 与基于Tucker分解的方法(MTuLR)相比提高了1.79、12.38和13.71个百分点。并在消融实验中证明了创新的有效性。

关键词

图像分类, 多分类, 低秩张量, 张量管道秩, 机器学习

Image Multi-Classification Model Based on Tensor Low Tubal Rank

Jiarui Zhang, Yuyu Hu, Kaiyu Tang, Yali Fan*

College of Science, University of Shanghai for Science and Technology, Shanghai

Received: Apr. 29th, 2024; accepted: May 23rd, 2024; published: May 31st, 2024

Abstract

Traditional machine learning methods often convert high-order tensor data into a lower-order

*通讯作者。

文章引用: 张家瑞, 胡毓榆, 唐开煜, 樊亚莉. 基于张量低管道秩的图像多分类模型[J]. 建模与仿真, 2024, 13(3): 3980-3997. DOI: 10.12677/mos.2024.133362

format for classification purposes. However, this approach can result in overfitting and the loss of tensor structure. To address these issues, this paper proposes a multi-classification model based on tensor low-tubal-rank (LRTMLR). The LRTMLR model directly classifies images in tensor format and utilizes the tensor pipe rank induced by tensor-tensor product and corresponding tensor kernel norm to handle low-rank tensors. This enables better utilization of the characteristics of the tensor structure and improves the accuracy of multi-classification for images in tensor format. The classification accuracy of LRTMLR was 9.6 percentage points higher than that of the methods without structure information (MLR) and with matrix structure information (LRMLR) on the three-class simulation data set, and 23.2 and 25.2 percentage points higher on the five-class simulation data set, respectively. On the three-class, five-class and 14-class subsets of the 101-class color image recognition dataset of Caltech University, the classification accuracy of LRTMLR was 10.01, 25.61 and 40.78 percentage points higher than that of MLR, and 10.68, 25.61 and 40.78 percentage points higher than that of LRMLR. Compared with the method based on CP decomposition (MCPLR), it is 6.47, 13.37 and 27.73 percentage points higher, and compared with the method based on Tucker decomposition (MTuLR), it is 1.79, 12.38 and 13.71 percentage points higher. The effectiveness of the innovation is demonstrated in ablation experiments.

Keywords

Image Classification, Multiple Classification, Low-Rank Tensor, Tensor Tubal Rank, Machine Learning

Copyright © 2024 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

分类问题是机器学习的核心问题之一，也是日常生活中最为常见的问题，在银行业务、网络安全、手写识别、互联网搜索以及图像处理等领域[1]都有着广泛的应用。许多传统的机器学习方法都可以用来解决多分类问题，如支持向量机[2]、k近邻法[3]、逻辑回归模型[4] [5]等。除此之外，近年来兴起的深度学习方法，如BP神经网络[6]和卷积神经网络[7]，也在分类问题上有着优异表现。虽然深度学习方法因其高分类精度为人所知，但其训练数据量大、参数学习耗时长、输入输出间不具有可解释性。在一些数据采集难度大，对分类器的可解释性要求较高的场合(如医疗影像数据分析)很难使用深度学习方法。因此本文依旧关注传统机器学习分类方法及其推广。

在众多分类问题中，图像分类作为计算机视觉的基本问题之一扮演着极其重要的角色。传统的机器学习算法大都基于向量，很难处理矩阵数据。但一些灰度图像以矩阵数据形式呈现，用基于向量的方法来处理会产生高维向量增加计算复杂度，也会损失矩阵数据的结构信息[8] [9]。因此，为了更好地利用结构信息，众多学者提出许多将向量模型推广到矩阵的方法。Gabriel等[10]基于矩阵图像数据的低秩特点，使用双线性回归模型代替传统的回归模型，提出了广义双线性回归模型(A Generalized Bilinear Regression Model, GBR)。但由于GBR压缩信息过多导致拟合误差太大，无法用于数据集建模，Hou等[11] [12]便在此基础上提出了多秩回归模型来进行矩阵图像数据分类。然而该模型是一个非凸优化问题，很难找寻全局解。于是受多秩模型思想构造低秩投影矩阵的启发，Yuan等[13]提出了低秩矩阵回归模型(Low-rank Matrix Regression, LRMR)，利用最小二乘损失加分类器的两步方法对矩阵数据进行分类。但是最小二乘

损失把图像数据直接和标签相减的处理方法在理论上并不合理,也不适用于一步分类,因此 Hu 等[14]提出了一般稳健低秩多分类逻辑回归(A General Robust Low-Rank Multinomial Logistic Regression, RLRMLR),利用交叉熵损失直接对属于每个类别的概率进行建模,使模型可以对矩阵数据一步分类,并且使用了图像恢复的思想使得模型可以对噪声数据稳健分类。

如今,大量的图像不以灰度图像的形式出现,更多以彩色图像或视频形式存在。这些彩色图像、视频和医学影像等都自然呈现张量形式。探索张量的内部结构信息,不仅可以用更少的参数更好的表示高维数据,同时也会使基于张量的模型更具有结构性和稳健性[15] [16],因此将传统分类算法扩展到其张量格式成为了近年来研究的热点。

支持向量机(Support Vector Machine, SVM)是传统分类算法中使用较广的一种方法,Tao 等[17]提出了一种监督张量学习(A Supervised Tensor Learning, STL)方案,将其扩展为支持张量机(Support Tensor Machine, STM)。支持张量机通过将权重参数建模为秩一张量,达到对张量分类的目的。但秩一张量表达信息的能力有限,因此 Kotsia 等[18] [19]基于 Tao 等的研究,采用了 CP 格式和 Tucker 格式取代 STM 中的秩一权重张量。除了支持向量机的张量扩展, Tan 等[20]还通过将权重张量建模为 CP 格式将逻辑回归模型推广到张量。然而,CP 秩定义为秩一张量分解的最小数,它的求解是一个 NP 难问题[21],Chen 等[22] [23]便使用了更具有扩展性的 TT 秩来推广支持张量机,使其能够更好的对高维张量数据进行分类。不过,以上方法都只提出了二分类的算法和实验,不适用于多分类数据。

文献[24]提出张量分类的难点在张量秩的选择与处理上。与矩阵秩不同,张量秩有多种不同的定义方式。最常见的是由 CP 分解诱导的 CP 秩和 Tucker 分解诱导的 Tucker 秩。对于低秩张量,利用结构信息的有效方式是对其张量秩加以约束。由于张量秩是非凸的,求解优化问题较为困难,于是普遍使用秩的凸近似即张量核范数对权重张量做低秩约束。CP 秩因其难以确定导致其凸近似难以处理,所以更好处理的 Tucker 秩[21]与其凸近似有着较为广泛的使用。然而,许多文献[25] [26] [27]也表明 Tucker 秩及其凸近似方法的使用是次优的,不能完全捕捉高维张量的结构相关性,因此, Lu 等[28]定义了一种基于张量-张量积的张量管道秩,并且提出了一种由张量-张量积严格诱导的张量核范数使它是张量管道秩的紧凸近似。该方法与矩阵方法有相似的性质和计算保证。Wang 等[29]在提出的模型中使用了 Lu 等提出的紧凸近似的相关定义和优化方法获得了不错的结果,进一步说明了张量管道秩方法的优越性。

但张量管道秩广泛应用于图像恢复问题,在图像分类上的应用并不多见。为了利用张量管道秩的优势解决图像多分类问题,本文受 Hu 等提出的多分类方法的启发,针对低秩张量的特点提出了低秩张量多分类逻辑回归(Low-Rank Tensor Multinomial Logistic Regression, LRTMLR)方法。这种方法使用了张量管道秩及其相应的张量核范数对低秩的权重张量做约束,最大程度利用张量的结构信息来提升张量数据的多分类精度。本文的主要工作和创新有以下三点:

- 1) 提出一种新的张量多分类模型,创新性地将多用于图像恢复领域的张量管道秩及其对应的核范数引入张量多分类模型作为惩罚项以更大程度利用张量结构信息;
- 2) 为了得到模型的最优参数,本文提出一种收敛的优化算法,引入辅助变量对模型进行解耦,利用张量奇异值分解以更新模型的张量核范数模块,并且证明了该优化算法的收敛性,保障实验的有效性;
- 3) 为验证本文方法的效果,在仿真数据和真实数据上进行大量实验,通过常用的多分类指标评价模型,结果表明在不同多分类情况下本文模型的分类精度较对比方法均取得明显提升。同时,为了验证引入核范数进行张量分类的有效性,设计了消融实验,结果表明分类类别越多,核范数惩罚项的贡献越大。

2. 预备知识

接下来将介绍张量及其运算中使用的符号和相关定义,并介绍 3 种常用的多分类方法。

2.1. 符号说明

本文将张量认为是向量(一阶张量)和矩阵(二阶张量)推广到高阶的多维数组, 用手写大写字母 $\mathcal{A} \in \mathbb{R}^{I_1 \times I_2 \times \dots \times I_d}$ 来表示 d 阶张量, 张量的每个元素表示为 $\mathcal{A}(i_1, i_2, \dots, i_d)$, 其中 $1 \leq i_k \leq I_k, k = 1, \dots, d$; 对于三阶张量 $\mathcal{B} \in \mathbb{R}^{I_1 \times I_2 \times I_3}$, 分别将其第 i 个前向切片、侧向切片和正向切片表示为 $\mathcal{B}(i, :, :)$ 、 $\mathcal{B}(:, i, :)$ 和 $\mathcal{B}(:, :, i)$, 用 $\mathcal{B}(i, j, :)$ 表示管道。用大写字母 $\mathbf{A}, \mathbf{B}, \dots$ 来表示矩阵; 用加粗的小写字母 $\mathbf{a}, \mathbf{b}, \dots$ 来表示向量; 用小写字母 a, b, \dots 来表示标量。同时, 本文使用 $\|\mathcal{A}\|_0$, $\|\mathcal{A}\|_1$, $\|\mathcal{A}\|_*$ 和 $\|\mathcal{A}\|_F$ 来表示 l_0 范数、 l_1 范数、核范数和 F 范数。

对于有 C 种可能标签的训练数据集 $\{\mathcal{X}_i \in \mathbb{R}^{I_1 \times I_2 \times \dots \times I_d} \mid i = 1, \dots, N\}$, 通常使用独热编码来表示标签 $\{\mathbf{y}_i \in \mathbb{R}^C \mid i = 1, \dots, N\}$ 。具体来说, 本文将 \mathcal{X}_i 的标签编码为 $\mathbf{y}_i = [0, \dots, 1, \dots, 0]^T \in \mathbb{R}^C$, 若 \mathcal{X}_i 属于第 j 类, 那么 \mathbf{y}_i 除了第 j 个元素为 1 外其余元素皆为 0。除此之外还定义 $\mathbf{Y} = [\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_N]^T \in \mathbb{R}^{N \times C}$ 为训练集标签矩阵。

2.2. 相关定义

定义 1 张量模 k 乘积。

对于一个张量 $\mathcal{A} \in \mathbb{R}^{I_1 \times I_2 \times \dots \times I_d}$ 与一个矩阵 $\mathbf{M} \in \mathbb{R}^{m \times k}$, 二者的模 k 乘积记作 $\mathcal{C} = \mathcal{A} \times_k \mathbf{M}$, 其定义如下:

$$\begin{aligned} \mathcal{C}(i_1, \dots, i_{k-1}, j, i_{k+1}, \dots, i_d) &= \mathcal{A} \times_k \mathbf{M} \\ &= \sum_{i_k=1}^{I_k} \mathbf{M}(j, i_k) \mathcal{A}(i_1, \dots, i_{k-1}, i_k, i_{k+1}, \dots, i_d). \end{aligned} \quad (1)$$

定义 2 张量内积。

对于张量 $\mathcal{A} \in \mathbb{R}^{I_1 \times I_2 \times \dots \times I_d}$ 和 $\mathcal{B} \in \mathbb{R}^{I_1 \times I_2 \times \dots \times I_d}$, 它们的内积 $\langle \mathcal{A}, \mathcal{B} \rangle$ 定义为

$$\langle \mathcal{A}, \mathcal{B} \rangle = \sum_{i_1=1}^{I_1} \sum_{i_2=1}^{I_2} \dots \sum_{i_d=1}^{I_d} \mathcal{A}(i_1, i_2, \dots, i_d) \mathcal{B}(i_1, i_2, \dots, i_d). \quad (2)$$

定义 3 张量的 Frobenius 范数。

张量 $\mathcal{A} \in \mathbb{R}^{I_1 \times I_2 \times \dots \times I_d}$ 的 Frobenius 范数定义为

$$\|\mathcal{A}\|_F = \sqrt{\langle \mathcal{A}, \mathcal{A} \rangle}. \quad (3)$$

定义 4 正交张量。

根据文献[30], 对于三阶张量 $\mathcal{A} \in \mathbb{R}^{I_1 \times I_2 \times I_3}$, 若它满足 $\mathcal{A}^* * \mathcal{A} = \mathcal{A} * \mathcal{A}^* = \mathcal{I}$, 则它为正交张量。

定义 5 张量 - 张量积(T 积)。

对于三阶张量, 设 $\mathcal{A} \in \mathbb{R}^{I_1 \times I_2 \times I_3}$, $\mathcal{B} \in \mathbb{R}^{I_2 \times m \times I_3}$, 根据文献[30], 有定义

$$\text{unfold}(\mathcal{A}) = \begin{bmatrix} \mathcal{A}(:, :, 1) \\ \mathcal{A}(:, :, 2) \\ \vdots \\ \mathcal{A}(:, :, I_3) \end{bmatrix} = \begin{bmatrix} \mathcal{A}^{(1)} \\ \mathcal{A}^{(3)} \\ \vdots \\ \mathcal{A}^{(I_3)} \end{bmatrix}. \quad (4)$$

$$\text{fold}(\text{unfold}(\mathcal{A})) = \mathcal{A}. \quad (5)$$

$$bcirc(\mathcal{A}) = \begin{bmatrix} \mathcal{A}^{(1)} & \mathcal{A}^{(I_3)} & \dots & \mathcal{A}^{(2)} \\ \mathcal{A}^{(2)} & \mathcal{A}^{(1)} & \dots & \mathcal{A}^{(3)} \\ \vdots & \vdots & \ddots & \vdots \\ \mathcal{A}^{(I_3)} & \mathcal{A}^{(I_3-1)} & \dots & \mathcal{A}^{(1)} \end{bmatrix}. \quad (6)$$

则二者的 T 积 $\mathcal{A} * \mathcal{B}$ 大小为 $I_1 \times m \times I_3$, 其定义为

$$\mathcal{A} * \mathcal{B} = fold(bcirc(\mathcal{A}) \cdot unfold(\mathcal{B})). \quad (7)$$

定义 6 张量的奇异值分解(T-SVD)。

设 $\mathcal{A} \in \mathbb{R}^{I_1 \times I_2 \times I_3}$, 根据文献[28], 它可以分解为

$$\mathcal{A} = \mathcal{U} * \mathcal{G} * \mathcal{V}^*. \quad (8)$$

其中 $\mathcal{U} \in \mathbb{R}^{I_1 \times I_1 \times I_3}$, $\mathcal{V} \in \mathbb{R}^{I_2 \times I_2 \times I_3}$ 是正交的, $\mathcal{G} \in \mathbb{R}^{I_1 \times I_2 \times I_3}$ 为每个正向切片都是对角矩阵的对角张量。

定义 7 张量管道秩。

根据 Mishia 等[31]和 Zhang 等[32]的研究, 对于张量 $\mathcal{A} \in \mathbb{R}^{I_1 \times I_2 \times I_3}$, 其张量管道秩记为 $rank_t(\mathcal{A})$, 定义为 \mathcal{G} 的非零的奇异管道的数量, 其中 \mathcal{G} 来源于 \mathcal{A} 的 T-SVD, 即 $\mathcal{A} = \mathcal{U} * \mathcal{G} * \mathcal{V}^*$, 于是有

$$rank_t(\mathcal{A}) = \#\{i, \mathcal{G}(i, i, :) \neq 0\}. \quad (9)$$

根据矩阵奇异值的递减性质和逆快速傅里叶变换, 文献[28]表明了 \mathcal{G} 的第一个正向切片 $\mathcal{G}(:, :, 1)$ 决定张量管道秩, 即

$$rank_t(\mathcal{A}) = \#\{i, \mathcal{G}(i, i, 1) \neq 0\}. \quad (10)$$

定义 8 张量核范数。

设 $\mathcal{A} = \mathcal{U} * \mathcal{G} * \mathcal{V}^*$ 是 $\mathcal{A} \in \mathbb{R}^{I_1 \times I_2 \times I_3}$ 的 T-SVD, 则根据文献[28], 张量 \mathcal{A} 的核范数定义为

$$\|\mathcal{A}\|_* := \langle \mathcal{G}, \mathcal{I} \rangle = \sum_{i=1}^r \mathcal{G}(i, i, 1). \quad (11)$$

其中 $r = rank_t(\mathcal{A})$, \mathcal{I} 为单位张量, 即第一个正向切片是单位矩阵其他正向切片都是零的张量。

2.3. 相关工作

接下来将分别介绍面向矩阵的无结构多分类模型、利用矩阵结构信息的多分类模型以及利用张量分解的多分类模型。

2.3.1. 无结构的多分类模型

在众多可以用来解决无结构多分类问题的方法中, 决策树模型对高维数据的微小变化较为敏感且容易过拟合, 而支持向量机对噪声较多的数据和样本重叠的情况表现不佳, 神经网络方法的可解释性较差, 并且神经网络方法的参数调整较为困难, 消耗计算资源较多, 而多分类逻辑回归模型(MLR)因其可解释性以及精确的分类结果成为处理无结构多分类数据最为热门的方法之一。

给定 C 类高维低秩张量训练数据集 $\{(\mathcal{X}_i, y_i), i=1, \dots, N\}$, 其中 $\{\mathcal{X}_i \in \mathbb{R}^{I_1 \times I_2 \times \dots \times I_d}\}$ 。将 \mathcal{X}_i 按照传统方式直接拉直成向量 $\{\mathbf{x}_i \in \mathbb{R}^{I_1 I_2 \dots I_d}\}$, 通过对 \mathbf{x}_i 属于第 r 类的条件概率进行建模, 其中条件概率如下:

$$P(Y_{ir} = 1 | \mathbf{x}_i; W, b) = \frac{e^{(\mathbf{w}^r)^T \mathbf{x}_i + b^r}}{\sum_{j=1}^C e^{(\mathbf{w}^j)^T \mathbf{x}_i + b^j}}. \quad (12)$$

其中 $W = [\mathbf{w}^1, \dots, \mathbf{w}^C] \in \mathbb{R}^{I_1 I_2 \dots I_d \times C}$, 使用最大似然估计得到多分类逻辑回归的目标函数

$$\min_{W,b} -\frac{1}{N} \sum_{i=1}^N \sum_{r=1}^C \log \left(\frac{e^{\langle w^r, x_i \rangle + b^r}}{\sum_{j=1}^C e^{\langle w^j, x_i \rangle + b^j}} \right). \quad (13)$$

其中 Y_{ir} 为 x_i 属于第 r 类的概率。由于式(13)是凸函数，我们可以使用梯度下降法轻易推导出全局解。

2.3.2. 带矩阵结构信息的多分类模型

上一节中的 MLR 算法是无结构多分类方法，没有利用数据中的结构信息，文献[14]提出的低秩多分类逻辑回归(Low-Rank Multinomial Logistic Regression, LRMLR)方法基于矩阵数据，利用了矩阵的结构信息，弥补了这一缺陷。

对于给定 C 类高维低秩张量训练数据集 $\{(X_i, y_i), i=1, \dots, N\}$ ，将 X_i 处理为矩阵形式，并使用矩阵核范数对低秩的权重张量做约束，以利用矩阵的结构信息，得到 LRMLR 的目标函数如下

$$\min_{W^r, b^r} F(X, W, b) + \lambda_1 \sum_{r=1}^C \|W^r\|_*. \quad (14)$$

其中

$$F(X, W, b) = -\frac{1}{N} \sum_{i=1}^N \sum_{r=1}^C Y_{ir} \log \left(\frac{e^{\langle w^r, X_i \rangle + b^r}}{\sum_{j=1}^C e^{\langle w^j, X_i \rangle + b^j}} \right). \quad (15)$$

但将张量处理为矩阵会破坏张量的结构，并且张量秩与矩阵秩大有不同。因此本文基于前人的工作，进一步利用张量结构信息，以提升高维张量数据的多分类效果。

2.3.3. 利用张量分解的多分类模型

上述两种方法将高阶张量展开为向量和矩阵形式，本质上依旧破坏了张量结构，损失了结构信息。因此张建光等[33]基于 CP 分解和 Tucker 分解提出的两种张量分类模型，对张量图像数据直接进行分类操作。

无论基于 CP 分解还是基于 Tucker 分解，其算法本质都是将权重张量进行分解，以此将无结构的分类模型扩展至相应的张量空间。文献[33]重点介绍了用于二分类的模型，同时也提出了推广至多分类的算法伪代码，根据伪代码我们得到基于 CP 分解和 Tucker 分解的多分类模型。

给定 C 类高维低秩张量训练数据集 $\{(X_i, y_i), i=1, \dots, N\}$ ，其中 $\{X_i \in \mathbb{R}^{I_1 \times I_2 \times \dots \times I_d}\}$ 。张量多分类逻辑回归模型可以表示为

$$\mathcal{F}(X, \mathcal{W}) = -\frac{1}{N} \sum_{i=1}^N \sum_{r=1}^C Y_{ir} \log \left(\frac{e^{\langle \mathcal{W}^r, X_i \rangle}}{\sum_{j=1}^C e^{\langle \mathcal{W}^j, X_i \rangle}} \right). \quad (16)$$

其中， $\mathcal{W}^r \in \mathbb{R}^{I_1 \times I_2 \times \dots \times I_d}$ ， $r=1, \dots, C$ 为权重张量。对式(16)的权重张量 \mathcal{W} 进行 CP 分解后，可以得到 N 个因子矩阵 $\{U^i\}_{i=1}^N$ ，即

$$\begin{aligned} \mathcal{W} &\approx \sum_{n=1}^R \mathcal{W}_n = \sum_{n=1}^R \mathbf{u}_n^1 \circ \mathbf{u}_n^2 \circ \dots \circ \mathbf{u}_n^N \\ &\triangleq \llbracket U^1, U^2, \dots, U^N \rrbracket. \end{aligned} \quad (17)$$

其中 R 是 CP 秩, \mathcal{W}_n 为第 n 个秩 1 张量, \mathbf{u}_n^i 为第 n 个秩 1 张量第 i 阶方向上的因子向量, \circ 为外积, $U^i = [\mathbf{u}_1^i, \mathbf{u}_2^i, \mathbf{u}_R^i]$ 是第 i 阶方向的因子矩阵。每一个因子矩阵列数相等, 且与初始张量秩相等。将式(17)代入式(16)则有

$$\mathcal{F}(\mathcal{X}) = -\frac{1}{N} \sum_{i=1}^N \sum_{r=1}^C Y_{ir} \log \left(\frac{e^{\langle \llbracket U^1, U^2, \dots, U^N \rrbracket, \mathcal{X}_i \rangle}}{\sum_{j=1}^C e^{\langle \llbracket U^1, U^2, \dots, U^N \rrbracket, \mathcal{X}_i \rangle}} \right). \quad (18)$$

式(18)有 N 个参数且对 N 个参数的非凸, 但若固定其余参数, 则式(18)对任意一个参数均为凸函数。因此在实验中我们将采用交替方向乘法(Alternating Direction Method of Multipliers, ADMM)对参数进行优化。

我们称上述基于 CP 分解的多分类模型为多分类 CP 逻辑回归模型(Multinomial CP Logistic Regression, MCPLR), 接下来将介绍基于 Tucker 分解的多分类 Tucker 逻辑回归模型(Multinomial Tucker Logistic Regression, MTuLR)。

Tucker 分解将权重张量 \mathcal{W} 分解为一个核张量与 N 阶方向上的因子矩阵的乘积, 即

$$\begin{aligned} \mathcal{W} &= \mathcal{M} \times_1 U_1 \times_2 U_2 \times_3 \dots \times_N U_N \\ &= \llbracket \mathcal{M}; U_1, U_2, \dots, U_N \rrbracket. \end{aligned} \quad (19)$$

其中 $\mathcal{M} \in \mathbb{R}^{R_1 \times R_2 \times \dots \times R_d}$ 为 Tucker 分解后的核张量, $\{U_i \in \mathbb{R}^{I_i \times R_i}\}_{i=1}^N$ 为沿 N 阶方向的因子矩阵, \times_i 表示核张量与因子矩阵的模 i 乘积, $R = [R_1, R_2, \dots, R_N]$ 为 Tucker 秩。将式(19)代入式(16)可得

$$\mathcal{F}(\mathcal{X}) = -\frac{1}{N} \sum_{i=1}^N \sum_{r=1}^C Y_{ir} \log \left(\frac{e^{\langle \llbracket \mathcal{M}; U_1, U_2, \dots, U_N \rrbracket, \mathcal{X}_i \rangle}}{\sum_{j=1}^C e^{\langle \llbracket \mathcal{M}; U_1^j, U_2^j, \dots, U_N^j \rrbracket, \mathcal{X}_i \rangle}} \right). \quad (20)$$

本文利用核范数进一步对张量结构信息加以利用, 使分类准确率较 MCPLR 和 MTuLR 有进一步提升。具体的对比实验结果将在第 5 节展示。

3. 低秩张量多分类模型(LRTMLR)

在高维张量数据中, 结构信息对分类结果起到至关重要的作用。利用张量结构信息可以提高模型表达能力、泛化能力和计算效率, 更有效捕捉数据中的空间时间等关联性, 缩减参数空间, 避免过拟合情况出现, 并增强模型可解释性, 有助于提升模型在各个领域中的性能和应用。

本文提出的低秩张量多分类逻辑回归(LRTMLR)方法通过对低秩权重张量的秩加以约束的方式充分考虑了张量结构, 与神经网络方法相比减少了运算成本, 并具有较强的可解释性。

给定 C 类张量训练数据集 $\{(\mathcal{X}_i, y_i), i=1, \dots, N\}$, LRTMLR 的目标函数为

$$\min_{\mathcal{W}^r, b^r} \mathcal{F}(\mathcal{X}, \mathcal{W}, b) + \lambda_1 \sum_{r=1}^C \|\mathcal{W}^r\|_*. \quad (21)$$

其中 $\mathcal{X}_i \in \mathbb{R}^{I_1 \times I_2 \times \dots \times I_d}$, $\mathcal{W}^r \in \mathbb{R}^{I_1 \times I_2 \times \dots \times I_d}$, $r=1, \dots, C$, \mathcal{W}^r 为属于第 r 类的张量所对应的权重张量, $\|\mathcal{W}^r\|_*$ 是张量管道秩对应的张量核范数, 代表权重张量的结构信息, $Y_{ir} \in \mathbb{R}^{N \times C}$ 为处理为独热格式的标签矩阵的 (i, r) 元, 以及 softmax 函数

$$\mathcal{F}(\mathcal{X}, \mathcal{W}, b) = -\frac{1}{N} \sum_{i=1}^N \sum_{r=1}^C Y_{ir} \log \left(\frac{e^{\langle \mathcal{W}^r, \mathcal{X}_i \rangle + b^r}}{\sum_{j=1}^C e^{\langle \mathcal{W}^j, \mathcal{X}_i \rangle + b^j}} \right). \quad (22)$$

由于张量秩的不唯一性，本文选择在同类问题中具有更好性能的张量管道秩来约束低秩权重张量 \mathcal{W}^r 。利用定义 5 和定义 6 中的 T 积和张量奇异值分解，按照定义 7 的方法诱导出 \mathcal{W}^r 的张量管道秩，并按照定义 8 的方式得到对应的张量核范数来处理 \mathcal{W}^r ，最终使模型可以利用到低秩张量的结构信息。

式(21)的第一项与第二项中同时出现了变量 \mathcal{W}^r ，使得目标函数的求解难度大大提升。为了降低求解难度，可以对模型进行解耦，增加约束条件使目标函数如下

$$\begin{aligned} \min_{\mathcal{W}^r, \mathcal{P}, b^r} \quad & \mathcal{F}(\mathcal{X}, \mathcal{W}, b) + \lambda_1 \sum_{r=1}^C \|\mathcal{P}^r\|_*, \\ \text{s.t} \quad & \mathcal{W}^r = \mathcal{P}^r. \end{aligned} \tag{23}$$

与文献[14]一致，本文使用 ADMM 算法来解上述优化问题，得到增广拉格朗日函数为

$$\begin{aligned} \mathcal{L}(\mathcal{X}, \mathcal{W}, b, \mathcal{P}, \mathcal{Q}, \mu) = & \mathcal{F}(\mathcal{X}, \mathcal{W}, b) \\ & + \lambda_1 \sum_{r=1}^C \|\mathcal{P}^r\|_* + \sum_{r=1}^C \langle \mathcal{Q}^r, \mathcal{P}^r - \mathcal{W}^r \rangle \\ & + \frac{\mu}{2} \sum_{r=1}^C \|\mathcal{P}^r - \mathcal{W}^r\|_F^2. \end{aligned} \tag{24}$$

为了方便表示，我们将张量堆叠成高一维度的张量，记作 $\mathcal{W} = \mathcal{W}_{stock}^r \in \mathbb{R}^{I_1 \times I_2 \times \dots \times I_d \times C}$ ， $\mathcal{P} = \mathcal{P}_{stock}^r \in \mathbb{R}^{I_1 \times I_2 \times \dots \times I_d \times C}$ ， $\mathcal{Q} = \mathcal{Q}_{stock}^r \in \mathbb{R}^{I_1 \times I_2 \times \dots \times I_d \times C}$ 为拉格朗日乘子张量， μ 和 λ_1 为惩罚参数。为了便于计算，本文采用了与 Yin 等[34]相同的方法，将惩罚参数取相同的初始值，并在迭代的每一步中更新。根据经验选择 10^{-3} 作为参数的初始值。

LRTMLR 算法的更新过程以 python 格式的伪代码形式表示为算法 1。

算法 1 低秩张量多项逻辑回归(LRTMLR)。

输入： $\{(\mathcal{X}, y_i), i=1, \dots, N\}, \lambda, \alpha$ 。

输出： \mathcal{W}, b 。

1) 初始化：

$\mathcal{W} = \text{np.zero}(\llbracket I_1, \dots, I_d, C \rrbracket)$ ，

$\mathcal{P} = \text{np.zero}(\llbracket I_1, \dots, I_d, C \rrbracket)$ ， $b = \text{np.zero}(\llbracket C, 1 \rrbracket)$ ，

$\mu = 1e^{-3}, \rho = 1.01$ ；

2) 更新 \mathcal{W}, b ：

for i in range(10):

$\mathcal{W} = \mathcal{W} - \alpha * \nabla_{\mathcal{W}} \mathcal{F}_{\mathcal{W}}$

$b = b - \alpha * \nabla_b \mathcal{F}_b$ ；

3) 更新 \mathcal{P} ：

for j in range(C):

 用张量奇异值收缩算法(TSVT)更新 \mathcal{P} ；

4) 更新 \mathcal{Q}, μ ：

$\mathcal{Q} = \mathcal{Q} + \mu * (\mathcal{P} - \mathcal{W})$ ；

$\mu = \rho * \mu$ ；

5) 输出 \mathcal{W}, b 。

其中 $\{\mathcal{P}^r\}_{r=1}^C$ 更新可通过文献[28]中提出的张量奇异值收缩算法(TSVT)来解决，本文将算法的基本流程写作算法 2。

算法 2 张量奇异值收缩算法(TSVT)。

输入: $\mathcal{X} \in \mathbb{R}^{I_1 \times I_2 \times I_3}, \tau > 0$.

输出: $\tilde{\mathcal{X}} = \mathcal{U} * \mathcal{G}_\tau * \mathcal{V}^*$.

1) 对 \mathcal{X} 的每个正向切片做快速傅里叶变换, 计算 $\bar{\mathcal{X}} = \text{np.fft.fft}(\mathcal{X}, \text{axis} = 2)$;

2) 在 $\bar{\mathcal{X}}$ 的每个正向切片上做矩阵 SVT:

for i in range($1, \lceil (I_3 + 1)/2 \rceil$):

$\mathcal{U}, \mathcal{G}, \mathcal{V} = \text{np.linalg.svd}(\bar{\mathcal{X}}[:, :, i])$;

$\bar{\mathcal{Z}}^{(i)} = \mathcal{U} \cdot (\mathcal{G} - \tau)_+ \cdot \mathcal{V}^*$.

for i in range($\lceil (I_3 + 1)/2 \rceil + 1, I_3$):

$\bar{\mathcal{Z}}[:, :, i] = \text{conj}(\bar{\mathcal{Z}}[:, :, I_3 - i - 3])$.

3) 对 $\bar{\mathcal{Z}}$ 的每个正向切片做逆快速傅里叶变换, 计算 $\tilde{\mathcal{X}} = \text{np.fft.ifft}(\bar{\mathcal{Z}}, \text{axis} = 2)$;

4) 输出 $\tilde{\mathcal{X}} = \mathcal{U} * \mathcal{G}_\tau * \mathcal{V}^*$ 。

4. 收敛性分析

Boyd 等[35]在研究中证明了双块 ADMM 算法的收敛性, 因此接下来本文通过说明本文提出的 LRTMLR 算法可以转化为双块标准 ADMM 来证明算法 1 的收敛性。

定理 1 算法 1 可转换为双块标准 ADMM。因此, 算法 1 是收敛的。

证明: 本文的优化问题如式(18)所示, 将其改写为标准形式如下

$$\begin{aligned} \min_{y_1, y_2} \quad & h(y_1) + \lambda_1 h(y_2); \\ \text{s.t.} \quad & A_1 y_1 + A_2 y_2 = z. \end{aligned} \quad (25)$$

其中各个参数的取值如式(26)所示:

$$\begin{cases} y_1 = [\mathbf{w}_1 \ \mathbf{w}_2 \ \cdots \ \mathbf{w}_C] \in \mathbb{R}^{I_1 I_2 \cdots I_d \times C}, \\ y_2 = [\mathbf{p}_1 \ \mathbf{p}_2 \ \cdots \ \mathbf{p}_C] \in \mathbb{R}^{I_1 I_2 \cdots I_d \times C}, \\ h(y_1) = \mathcal{F}(\mathcal{X}, \mathcal{W}, b), \\ h(y_2) = \sum_{r=1}^C \|\mathcal{P}^r\|_*, \\ \mathbf{w}_r = \text{vec}(\mathcal{W}^r), \\ \mathbf{p}_r = \text{vec}(\mathcal{P}^r), \\ A_1 = [\mathbf{e} \ \mathbf{e} \ \cdots \ \mathbf{e}]^T \in \mathbb{R}^{C \times I_1 I_2 \cdots I_d}, \\ A_2 = [-\mathbf{e} \ -\mathbf{e} \ \cdots \ -\mathbf{e}]^T \in \mathbb{R}^{C \times I_1 I_2 \cdots I_d}, \\ z = [\mathbf{0} \ \mathbf{0} \ \cdots \ \mathbf{0}]^T \in \mathbb{R}^{C \times I_1 I_2 \cdots I_d}. \end{cases} \quad (26)$$

如式(22)和(26)所示, $h(y_1)$ 为 softmaximum 函数, 它是 hardmaximum 函数的近似, 并且同 hard 一样是凸函数。如式(21)所示, $h(y_2)$ 是张量核范数的加和, 而核范数是张量奇异值的和, 作为张量秩的凸优化被提出, 因此 $h(y_1)$ 和 $h(y_2)$ 皆为对于参数的凸函数, 优化问题式(23)可转化为如式(25)所示的标准双块 ADMM, 满足 ADMM 算法收敛的假设, 故算法 1 收敛得证。

5. 实验

接下来我们将对提出的模型进行检验, 分别使用不同方法在仿真数据集和真实数据集上进行分类实

验, 并以不同的评价指标来评估本文方法和相关工作中提到的方法在分类方面的优劣。最后, 为了验证本文创新的有效性, 本文设计了消融实验以探索模型不同模块对分类效果的影响。

实验采用交叉验证方法, 使用网格搜索方法来进行参数选择。为制表规整, 使用 Lbd1、Lbd2、Lbd3 指代模型的惩罚参数 λ_1 、 λ_2 、 λ_3 , 以 Alpha 和 Gamma 指代执行 ADMM 算法时用到的步长 α 和 γ 。在进行网格搜索选择参数时, 根据经验决定 $\{\lambda_i | i=1,2,3\}$ 的取值范围为 $[1e-5, 1e3]$, α 与 γ 的取值范围为 $[0.1, 1.5]$, CP 秩的初始值为 3, Tucker 秩的初始值为 $[3, 3, 3, 3]$ 。

本章所有实验均在搭载 2.50GHz i5-12500H 处理器 64 位操作系统的计算机上使用 python3.7 软件实现。

5.1. 评价指标

评价多分类模型的指标有两种类型。一类是由多个二分类评价指标转化而来的多分类评价指标, 另一类是直接用于多分类的指标。

根据文献[1], 对于二分类问题, 准确率(Accuracy)、精确率(Precision)、召回率(Recall)和 F1-分数(F1-Score)是常用的评价指标。通常以关注的类为正类, 其他类别为负类, 准确率指正类和负类中预测正确的数量占总量的比例, 精确率指在预测为正类的样本中真正类所占的比例, 召回率指在所有的正类中被预测为正类的比例, F1-分数为精确率和召回率的调和平均数。F1-分数扩展到多分类的两个指标为宏观 F1 (Macro-F1)和微观 F1 (Micro-F1)。前者是对每个标签分别计算 F1-分数再取不加权平均的评价指标, 后者则是将多个二分类评价对应相加计算精确率和召回率最后求得 F1-分数, 宏观 F1 和微观 F1 越高, 模型越适合。

Kappa 系数[36]和海明距离(Hamming Distance) [37]作为直接评价多分类模型的指标也被广泛使用。Kappa 系数是一种用在统计学中评估一致性的方法, 值越高代表模型实现的分类准确度越高, 反之则代表分类准确度越低。海明距离可以衡量预测标签与真实标签之间的距离, 距离越小表示分类效果越好。

综合各种评价指标的特点, 本文最终采用准确率、Kappa 系数、海明距离、宏观 F1 和微观来评价模型。

5.2. 仿真实验

为了说明本文模型的分类效果, 我们设计了仿真实验, 将本文提出的利用了低秩张量结构信息的张量多分类模型(LRTMLR)与无结构的多分类模型(MLR)、带矩阵结构信息的多分类模型(LRMLR)作比较, 通过各项多分类评价指标来反应模型的有效性。

在均匀分布下随机生成大小为 $n1 \times r \times n3$ 和 $r \times n2 \times n3$ 的两个临时张量, 对二者按照定义 5 中的 t 积做运算, 得到所需的大小为 $n1 \times n2 \times n3$ 的目标张量。其中 $n1$ 、 $n2$ 、 $n3$ 为目标张量大小, r 为张量管道秩的大小, 固定生成大小为 $32 \times 32 \times 3$ 、管道秩为 3 的目标张量, 重复多次, 同时生成对应的标签构成数据量为 500 和 1000 的三分类和五分类数据集。

由于 MLR 和 LRMLR 模型并不能直接处理张量数据, 本文在应用这两个模型时对数据进行了额外的预处理。对于 MLR 模型, 我们直接将张量数据拉直成向量, 即 $\mathbf{x}_i \in \mathbb{R}^{3072}$ 。对于 LRMLR 模型, 为了将张量数据变为矩阵数据应用模型, 本文参考 python 官方手册, 采用了更常用也更适用于矩阵模型的浮点算法, 将张量数据的正向切片按照一定比重做加法合并为一个矩阵。对于彩色图片可以按照 RGB 通道将其分开为 R 、 G 、 B 三个矩阵, 再以 $X_i = 0.299R + 0.587G + 0.114B$ 公式转换为相应的灰度图片, 使其数据格式变为 $X_i \in \mathbb{R}^{32 \times 32}$ 。仿真数据和真实数据均使用这种方式来将张量数据处理为矩阵。

5.2.1. 三分类实验

三分类数据集包含 500 个三阶张量, 随机划分训练集和测试集, 用其中 250 个数据进行训练, 另外

250 个进行测试，经过相应的预处理之后分别计算 Acc、MacroF1、MicroF1、Kappa 系数和 Hamm 距离，结果如图 1 所示，可以看出在不同评价指标下 LRTMLR 都比其他两种方法的效果好。

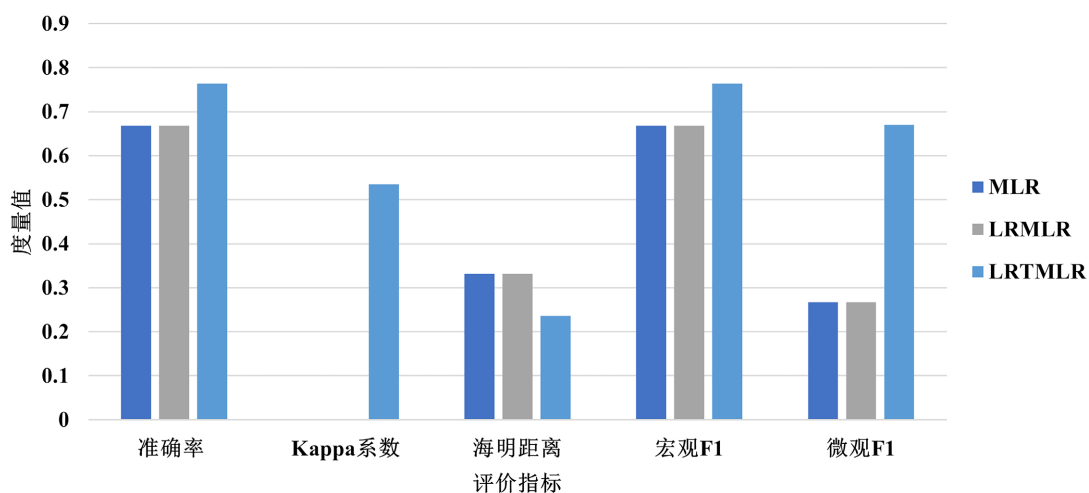


Figure 1. Results of 3-class experiments on simulated data
图 1. 仿真数据三分类实验结果

5.2.2. 五分类实验

五分类数据集包含 1000 个三阶张量，随机划分训练集和测试集，用其中 500 个数据进行训练，其余 500 个做测试，得到五种评价指标的结果如图 2 所示，可以看出在不同评价指标下 LRTMLR 都具有优势。

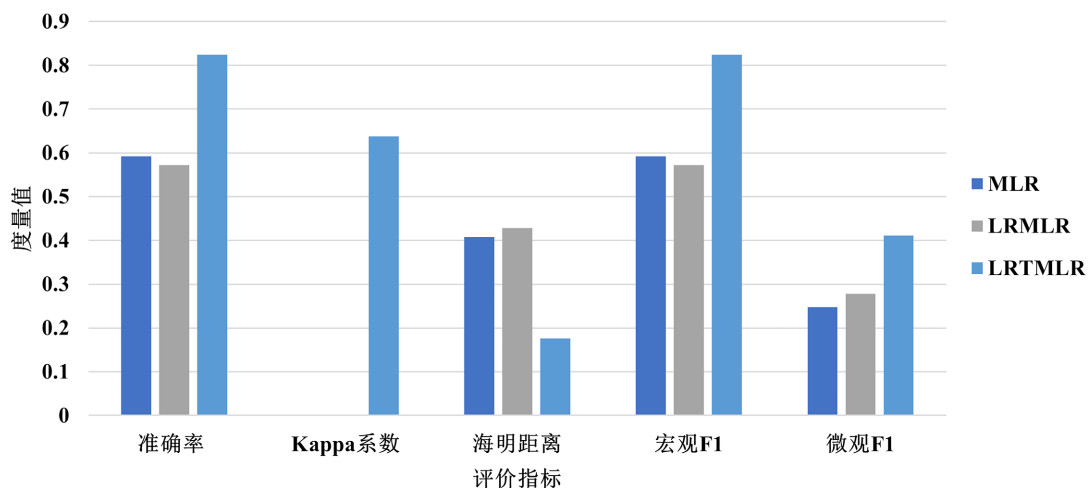


Figure 2. Results of 5-class experiments on simulated data
图 2. 仿真数据五分类实验结果

在仿真数据上的实验结果反应了本文方法的分类优势，接下来将使用真实数据集进一步验证模型在应用时的表现情况。

5.3. 真实数据

本文选择加州理工大学的 101 类彩色图像识别数据集来进行真实数据的实验，该数据集总共有 102

类 9145 幅彩色图像，其中包括一个背景类即杂乱类。该数据集示例如图 3 所示。在本文实验中，分别取数据集包含不同数据量的子集进行三次多分类实验。三个子数据集分别包含图片数据量为 896 张、1406 张和 1940 张，对应三分类、五分类和十四分类的实验。为了方便实验，统一将数据格式重塑为 $32 \times 32 \times 3$ ，即 $\mathcal{X}_i \in \mathbb{R}^{32 \times 32 \times 3}$ 。向量方法和矩阵方法的预处理与仿真实验中一致。



Figure 3. Examples of real datasets
图 3. 真实数据集示例

接下来将对本文方法 LRTMLR 和 MLR、LRMLR、MCPLR、MTuLR 四种方法的分类效果，其中 MLR 和 LRMLR 为破坏了张量结构的无结构和带矩阵结构信息的多分类方法，MCPLR 和 MTuLR 为没有破坏张量结构的、基于 CP 分解和 Tucker 分解的多分类方法。LRTMLR 与 MLR、LRMLR 的对比实验是为了体现以张量格式输入的优势，与 MCPLR、MTuLR 的对比实验是为了证明应用核范数利用张量结构信息对分类效果的提升，由于对比实验的目的不同，因此在下述不同类别的实验中分为两组描述。

5.3.1. 三分类实验

用于三分类实验的数据集总共包含了 896 张彩色图片，用其中 448 张图片进行训练，另外 448 张进行测试。由于对比维度的不同，首先介绍 LRTMLR 与 MLR、LRMLR 的对比实验。我们通过大量实验得到三个模型各自的最优结果，所选择的参数如表 1 所示。

Table 1. Parameter setting of 3-class experiment on real data (destroy/not destroy tensor structure)
表 1. 真实数据三分类实验参数设置(破坏/不破坏张量结构)

	Alpha	Gamma	Lbd1	Lbd2	Lbd3
MLR	0.1	/	/	/	/
LRMLR	0.5	1.2	1e-1	1e-3	1e-3
LRTMLR	1.2	/	1e-5	/	/

实验得到五种评价指标的结果如表 2 所示,可以看出在不同评价指标下 LRTMLR 都具有相当的优势。其中, LRTMLR 的分类准确率与 MLR 相比提升了 6.47 个百分点,与 LRMLR 相比提升了 1.79 个百分点。

Table 2. Results of 3-class experiments on real data (destroy/not destroy tensor structure)
表 2. 真实数据三分类实验结果(破坏/不破坏张量结构)

	准确率	Kappa 系数	海明距离	宏观 F1	微观 F1
MLR	0.8996	0.0000	0.1004	0.8996	0.3157
LRMLR	0.8929	0.0087	0.1071	0.8928	0.3145
LRTMLR	0.9997	0.8730	0.0223	0.9777	0.8809

接下来将介绍 LRTMLR 与 MCPLR、MTuLR 的对比实验结果。如表 3 所示,可以看出在各个评价指标下 LRTMLR 都优于其余两种方法。与 MCPLR 相比, LRTMLR 的分类准确率提升了 6.47 个百分点,与 MTuLR 相比, LRTMLR 的分类准确率提升了 1.79 个百分点。

Table 3. Results of 3-class experiments on real data (Whether to use the nuclear norm)
表 3. 真实数据三分类实验结果(是否使用核范数)

	准确率	Kappa 系数	海明距离	宏观 F1	微观 F1
MCPLR	0.9107	0.6212	0.0893	0.7613	0.9107
MTuLR	0.9576	0.7619	0.0424	0.8230	0.9576
LRTMLR	0.9754	0.8556	0.0246	0.8558	0.9754

根据上述实验结果,证明本文方法在三分类彩色图片数据集上的优势。

5.3.2. 五分类实验

用于五分类实验的数据集总共包含了 1406 张彩色图片,同样将图片格式经过相应预处理应用于三个模型,用其中 703 张图片进行训练,其余的进行测试。首先介绍 LRTMLR 与 MLR、LRMLR 的对比实验,所选择的参数如表 4 所示。

Table 4. Parameter setting of 5-class experiment on real data (destroy/not destroy tensor structure)
表 4. 真实数据五分类实验参数设置(破坏/不破坏张量结构)

	Alpha	Gamma	Lbd1	Lbd2	Lbd3
MLR	0.1	/	/	/	/
LRMLR	0.3	0.3	1e-1	1e-5	1e-3
LRTMLR	0.5	/	1e-4	/	/

实验得到五种评价指标的结果如表 5 所示,可以看出在不同评价指标下 LRTMLR 都具有优势。其中 LRTMLR 的五分类准确率较 MLR 提升了 25.61 个百分点,较 LRMLR 提升了 27.46 个百分点,分类效果显著提升。

Table 5. Results of 5-class experiments on real data (destroy/not destroy tensor structure)**表 5.** 真实数据五分类实验结果(破坏/不破坏张量结构)

	准确率	Kappa 系数	海明距离	宏观 F1	微观 F1
MLR	0.5775	0.0000	0.4225	0.5775	0.1464
LRMLR	0.5590	0.0338	0.4410	0.5590	0.1818
LRTMLR	0.8336	0.6898	0.1664	0.8335	0.5994

接下来将介绍 LRTMLR 与 MCPLR、MTuLR 的对比实验结果。如表 6 所示,可以看出在各个评价指标下 LRTMLR 都优于其余两种方法。与 MCPLR 相比,LRTMLR 的分类准确率提升了 13.37 个百分点,与 MTuLR 相比则提升了 12.38 个百分点,本文方法的分类效果明显好于其它两种方法。

Table 6. Results of 5-class experiments on real data (whether to use the nuclear norm)**表 6.** 真实数据五分类实验结果(是否使用核范数)

	准确率	Kappa 系数	海明距离	宏观 F1	微观 F1
MCPLR	0.6984	0.4797	0.3016	0.3210	0.6984
MTuLR	0.7084	0.5013	0.2916	0.4001	0.7084
LRTMLR	0.8321	0.6878	0.1679	0.5955	0.8321

根据上述实验结果,证明本文方法在五分类彩色图片数据集上的优势。

5.3.3. 十四分类实验

用于十四分类实验的数据集总共包含了 1940 张彩色图片,用其中 970 张图片进行训练,其余的进行测试。首先介绍 LRTMLR 与 MLR、LRMLR 的对比实验,所选择的参数如表 7 所示。

Table 7. Parameter setting of 14-class experiment on real data (destroy/not destroy tensor structure)**表 7.** 真实数据十四分类实验参数设置(破坏/不破坏张量结构)

	Alpha	Gamma	Lbd1	Lbd2	Lbd3
MLR	0.1	/	/	/	/
LRMLR	1	0.7	1e-1	1e-4	1e-5
LRTMLR	0.1	/	1e-3	/	/

实验得到五种评价指标的结果如表 8 所示,尽管 LRTMLR 在海明距离上的表现不如 MLR 和 LRMLR,但它的分类准确率却较其他两种方法均提升了 40.78 个百分点。从这些分类结果可以看出,在向量和矩阵方法的准确率表现极为不佳的情况下,张量方法 LRTMLR 依然能保持分类精度方面的相对优势。

接下来将介绍 LRTMLR 与 MCPLR、MtuLR 的对比实验结果。如表 9 所示,可以看到与 MCPLR 相比,LRTMLR 的分类准确率提升了 27.73 个百分点,与 MtuLR 相比则提升了 13.71 个百分点。

通过对比分类准确率提升的百分点,可以看出训练数据集类别越多,本文方法的效果较其余方法更具有更多优势,这是因为本文的方法充分考虑到张量数据本身的结构信息。

Table 8. Results of 14-class experiments on real data (destroy/not destroy tensor structure)
表 8. 真实数据十四分类实验结果(破坏/不破坏张量结构)

	准确率	Kappa 系数	海明距离	宏观 F1	微观 F1
MLR	0.4258	0.0000	0.5742	0.4258	0.0459
LRMLR	0.4258	0.0000	0.5742	0.8928	0.0459
LRTMLR	0.8336	0.2556	0.6051	0.9777	0.3132

Table 9. Results of 14-class experiments on real data (whether to use the nuclear norm)
表 9. 真实数据十四分类实验结果(是否使用核范数)

	准确率	Kappa 系数	海明距离	宏观 F1	微观 F1
MCPLR	0.3278	0.2000	0.6722	0.0775	0.3278
MtuLR	0.4680	0.2145	0.5320	0.1065	0.4680
LRTMLR	0.6052	0.4556	0.3948	0.3132	0.6052

5.3.4. 时间消耗

为了比较各个方法的复杂度，我们在 python 上做单次实验并记录消耗的时间。其中 MLR、LRMLR 和 LRTMLR 方法的消耗时间如图 4 所示，MCPLR、MtuLR 和 LRTMLR 的消耗时间如图 5 所示。

通过图 4 可以看出，向量方法因其完全忽略图像结构复杂度低所耗费时间也最少，利用了部分结构信息的矩阵方法的耗时几乎是张量方法的十倍，该结果也符合理论预期。通过图 5 可以看到，LRTMLR 在三分类和五分类下比 MCPLR 的消耗时间少，在十四分类实验下比 MCPLR 仅多不到 2.3 s。而 MtuLR 因其将权重张量分解为一个核心张量和因子矩阵，更具有结构化，耗时较其余两种方法更短，结果符合预期。考虑到 LRTMLR 的分类精度，同等情况下本文方法依旧具有优势。

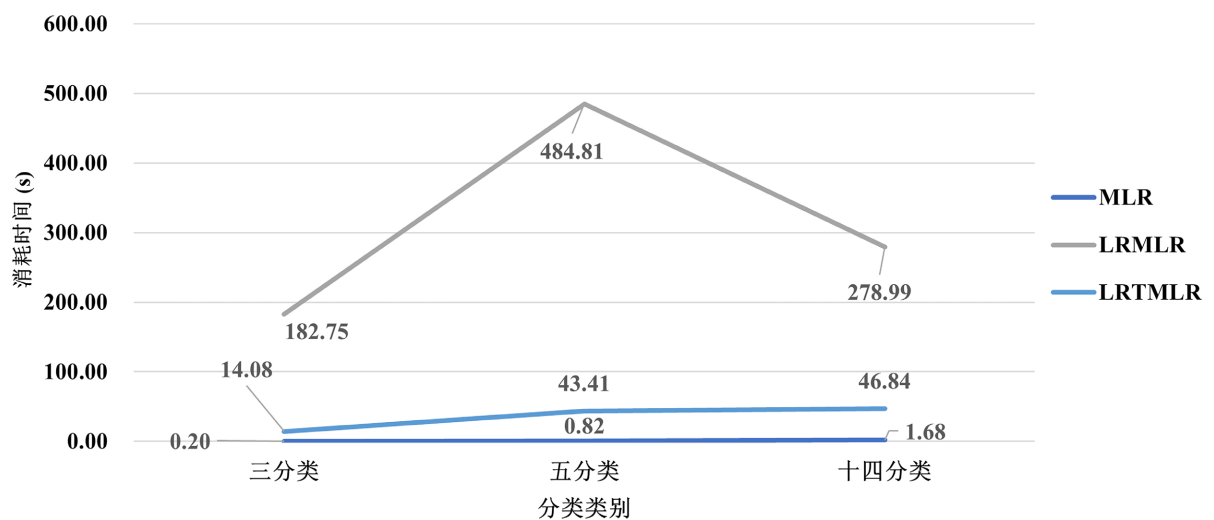


Figure 4. CPU times for different datasets (destroy/not destroy tensor structure)
图 4. 不同数据集的 CPU 时间(破坏/不破坏张量结构)

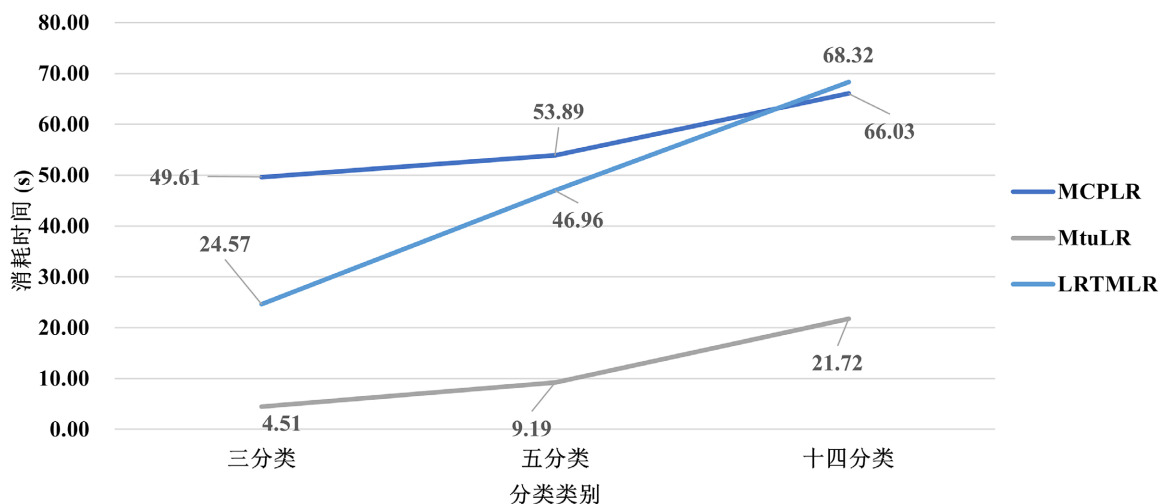


Figure 5. CPU times for different datasets (Whether to use the nuclear norm)

图 5. 不同数据集的 CPU 时间(是否使用核范数)

5.3.5. 消融实验

为进一步评估模型中两个模块对分类效果的影响，本文设计了消融实验，接下来将介绍消融实验结果。将 LRTMLR 模型分为 h_1 和 h_2 两个模块，即 $h_1 = \mathcal{F}(\mathcal{X}, \mathcal{W}, b)$ 为交叉熵损失项， $h_2 = \sum_{r=1}^C \|\mathcal{W}^r\|_*$ 为核范数惩罚项。当模型仅包含 h_1 时，模型便退化到无结构多分类模型 MLR。但消融实验与 5.3 节 LRTMLR 和 MLR 的对比实验的区别在于，5.3 节在进行 MLR 实验时按照主流方法将输入展开为向量，而消融仅是为了探索两个模块对 LRTMLR 分类结果的贡献，仅包含 h_1 的实验输入依旧使用张量格式。

分别在移除 h_2 仅包含 h_1 和同时包含 $h_1 + h_2$ 的条件下进行实验，为方便起见，在表格中我们用 h1 代替 h_1 ，用 h1 + h2 代替 $h_1 + h_2$ ，得到的结果如下表 10 所示。

可以看到，尽管在三分类数据集上， $h_1 + h_2$ 的宏观 F1 表现不足，但在其余四项指标下都较 h_1 有所提升。并且随着分类类别的增多，核范数惩罚项 h_2 对模型的影响越发明显， $h_1 + h_2$ 的分类效果较仅包含交叉熵损失的 h_1 有了明显提升，在三分类、五分类、十四分类数据集下， $h_1 + h_2$ 的分类准确率分别较仅包含交叉熵损失的 h_1 提升了 0.45、10.24、17.84 个百分点。以此证明了本文创新性地加入核范数惩罚项是有效的。

Table 10. Ablation experimental results on different categories of data sets

表 10. 不同类别数据集上消融实验结果

	三分类		五分类		十四分类	
	h1	h1 + h2	h1	h1 + h2	h1	h1 + h2
准确率	0.9710	0.9754	0.7297	0.8321	0.4268	0.6052
kappa 系数	0.8402	0.8556	0.5094	0.6878	0.2491	0.4556
海明距离	0.0290	0.0246	0.2703	0.1679	0.5732	0.3948
微观 F1	0.9710	0.9754	0.7297	0.8321	0.4268	0.6052
宏观 F1	0.8581	0.8558	0.5228	0.5955	0.2021	0.3132
消耗时间	1.3531	15.4131	3.5260	32.6684	6.7978	50.0029

6. 结语

本文基于利用矩阵结构信息的多分类逻辑回归,提出了一种能够直接处理张量格式数据、利用张量结构信息的新的多分类模型。该模型使用张量管道秩及其对应的张量核范数,通过解决一个秩最小化的问题充分利用了低秩张量的结构信息,对比无结构的多分类模型、带矩阵结构信息的多分类模型和基于CP分解、Tucker分解的带张量结构信息的多分类模型,在张量数据分类效果上取得了显著优势,也为张量多分类提出了新的思路。

消融实验表明,本文将张量管道秩及其核范数作为惩罚项加入目标函数以利用张量的结构信息的创新显著提升了分类效果与模型性能。模型面向的多分类数据集类别越多,核范数发挥的作用越大,有效保证了分类的准确率。

考虑到张量管道秩及其张量核范数在图像恢复领域的显著优势,以及现实生活中大量可采集的张量图像都是有噪声的,若是直接使用带噪声的数据进行分类会影响分类器的准确度。因此,为分类器增加稳健项使其可以稳健地对有噪声张量数据进行分类将成为下一步的研究方向,同时,考虑到张量管道秩只适用于三阶张量,对于更高阶的张量,需要应用不同的张量秩及其相应的核范数对模型进行进一步优化和推广。

基金项目

国家自然科学基金资助项目(12371308)。

参考文献

- [1] 李航. 统计学习方法[M]. 第2版. 北京: 清华大学出版社, 2019.
- [2] Corinna, C. and Vladimir, N.V. (1995) Support-Vector Networks. *Machine Learning*, **20**, 273-297. <https://doi.org/10.1007/BF00994018>
- [3] Gregory, S., Trevor, D. and Piotr, I. (2008) Nearest-Neighbor Methods in Learning and Vision. *IEEE Transactions on Neural Networks*, **19**, 377. <https://doi.org/10.1109/TNN.2008.917504>
- [4] Chanyeong, K. and Alan, C.-M. (2002) Multinomial Logistic Regression. *Nursing Research*, **51**, 404-410. <https://doi.org/10.1097/00006199-200211000-00009>
- [5] Böhning, D. (1992) Multinomial Logistic Regression Algorithm. *Annals of the Institute of Statistical Mathematics*, **44**, 197-200. <https://doi.org/10.1007/BF00048682>
- [6] Erb, R.J. (1993) Introduction to Backpropagation Neural Network Computation. *Pharmaceutical Research*, **10**, 165-170. <https://doi.org/10.1023/A:1018966222807>
- [7] Neelapu, R., Devi, G.L. and Rao, K.S. (2018) Deep Learning Based Conventional Neural Network Architecture for Medical Image Classification. *Traitement du Signal*, **35**, 169-182.
- [8] Song, K., Nie, F.P., Han, J.W., et al. (2017) Parameter Free Large Margin nearest Neighbor for Distance Metric Learning. *Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence*, San Francisco, 4-9 February 2017, 2555-2561. <https://doi.org/10.1609/aaai.v31i1.10861>
- [9] Cai, D., He, X.F., Hu, Y.X., et al. (2007) Learning a Spatially Smooth Subspace for Face Recognition. *Proceedings of the 2007 IEEE Conference on Computer Vision and Pattern Recognition*, Minneapolis, 17-22 June 2007, 1-7. <https://doi.org/10.1109/CVPR.2007.383054>
- [10] Gabriel, K.R. (1998) Generalised Bilinear Regression. *Biometrika*, **85**, 689-700. <https://doi.org/10.1093/biomet/85.3.689>
- [11] Hou, C.P., Nie, F.P., Yi, D. and Wu, Y. (2012) Efficient Image Classification via Multiple Rank Regression. *IEEE Transactions on Image Processing*, **22**, 340-352. <https://doi.org/10.1109/TIP.2012.2214044>
- [12] Hou, C.P., Jiao, Y.Y., Nie, F.P., et al. (2017) 2D Feature Selection by Sparse Matrix Regression. *IEEE Transactions on Image Processing*, **26**, 4255-4268. <https://doi.org/10.1109/TIP.2017.2713948>
- [13] Yuan, H.L., Li, J.Y., Loi, L.L., et al. (2020) Low-Rank Matrix Regression for Image Feature Extraction and Feature Selection. *Information Sciences*, **522**, 214-226. <https://doi.org/10.1016/j.ins.2020.02.070>

- [14] Hu, Y.Y., Fan, Y.L., Song, Y., *et al.* (2023) A General Robust Low-Rank Multinomial Logistic Regression for Corrupted Matrix Data Classification. *Applied Intelligence*, **53**, 18564-18580. <https://doi.org/10.1007/s10489-022-04424-0>
- [15] Liu, J.N., Zhu, C., Long, Z., *et al.* (2021) Low-Rank Tensor Ring Learning for Multi-Linear Regression. *Pattern Recognition*, **113**, Article ID: 107753. <https://doi.org/10.1016/j.patcog.2020.107753>
- [16] Koniusz, P., Wang, L. and Cherian, A. (2022) Tensor Representations for Action Recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **44**, 648-665. <https://doi.org/10.1109/TPAMI.2021.3107160>
- [17] Tao, D.C., Li, X.L., Hu, W.M., *et al.* (2005) Supervised Tensor Learning. *Proceedings of the 5th IEEE International Conference on Data Mining (ICDM'05)*, Houston, 27-30 November 2005, 8.
- [18] Kotsia, I., Guo, W.W. and Patras, I. (2012) Higher Rank Support Tensor Machines for Visual Recognition. *Pattern Recognition*, **45**, 4192-4203. <https://doi.org/10.1016/j.patcog.2012.04.033>
- [19] Kotsia, I. and Patras, I. (2011) Support Tucker Machines. *Proceedings of the 24th IEEE Conference on Computer Vision and Pattern Recognition*, Colorado Springs, 20-25 June 2011, 633-640.
- [20] Tan, X., Zhang, Y., Tang, S.L., *et al.* (2012) Logistic Tensor Regression for Classification. *Proceedings of the Third Sino-Foreign-Interchange Conference on Intelligent Science and Intelligent Data Engineering*, Nanjing, 15-17 October 2012, 589-597.
- [21] Kolda, T.G. and Bader, B.W. (2009) Tensor Decompositions and Applications. *Society for Industrial and Applied Mathematics*, **51**, 455-500. <https://doi.org/10.1137/07070111X>
- [22] Chen, C., Batselier, K., Ko, C.-Y., *et al.* (2018) A Support Tensor Train Machine. *Proceedings of the 2019 International Joint Conference on Neural Networks*, Budapest, 14-19 July 2019, 1-8. <https://doi.org/10.1109/IJCNN.2019.8851985>
- [23] Chen, C., Batselier, K., Yu, W.J., *et al.* (2022) Kernelized Support Tensor Train Machines. *Pattern Recognition*, **122**, Article ID: 108337. <https://doi.org/10.1016/j.patcog.2021.108337>
- [24] Yang, J.-H., Zhao, X.-L., Ji, T.-Y., *et al.* (2020) Low-Rank Tensor Train for Tensor Robust Principal Component Analysis. *Applied Mathematics and Computation*, **367**, Article ID: 124783. <https://doi.org/10.1016/j.amc.2019.124783>
- [25] Liu, J., Musialski, P., Wonka, P., *et al.* (2013) Tensor Completion for Estimating Missing Values in Visual Data. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **35**, 208-220. <https://doi.org/10.1109/TPAMI.2012.39>
- [26] Gandy, S., Recht, B. and Yamada, I. (2011) Tensor Completion and Low-rank Tensor Recovery via Convex Optimization. *Inverse Problems*, **27**, Article ID: 025010. <https://doi.org/10.1088/0266-5611/27/2/025010>
- [27] Tomioka, R., Hayashi, K. and Kashima, H. (2010) Estimation of Low-Rank Tensors via Convex Optimization. arXiv:1010.0789
- [28] Lu, C.Y., Feng, J.S., Liu, W., *et al.* (2020) Tensor Robust Principal Component Analysis with a New Tensor Nuclear Norm. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **42**, 925-938. <https://doi.org/10.1109/TPAMI.2019.2891760>
- [29] Wang, H.L., Zhang, F., Wang, J.Y., *et al.* (2020) Estimating Structural Missing Values via Low-Tubal-Rank Tensor Completion. 2020-2020 *IEEE International Conference on Acoustics, Speech and Signal Processing*, Barcelona, 4-8 May 2020, 3297-3330. <https://doi.org/10.1109/ICASSP40776.2020.9054531>
- [30] Kilmer, M.E. and Martin, C.D. (2011) Factorization Strategies for Third-Order Tensors. *Linear Algebra and Its Applications*, **435**, 641-695. <https://doi.org/10.1016/j.laa.2010.09.020>
- [31] Kilmer, M.E., Braman, K., Hao, N. and Hoover, R.C. (2013) Third-Order Tensors as Operators on Matrices: A Theoretical and Computational Framework with Applications in Imaging. *SIAM Journal on Matrix Analysis and Applications*, **34**, 148-172. <https://doi.org/10.1137/110837711>
- [32] Zhang, Z.M., Ely, G., Aeron, S., *et al.* (2014) Novel Methods for Multilinear Data Completion and De-Noising Based on Tensor-SVD. *Proceedings of the 2014 IEEE Conference on Computer Vision and Pattern Recognition*, Columbus, 23-28 June 2014, 3842-3849. <https://doi.org/10.1109/CVPR.2014.485>
- [33] 张建光, 韩亚洪. 基于张量分解的逻辑回归图像分类算法[EB/OL]. <http://www.paper.edu.cn/releasepaper/content/201605-772>, 2016-05-20.
- [34] Yin, M., Zeng, D.Y., Gao, J.B., *et al.* (2018) Robust Multinomial Logistic Regression Based on RPCA. *IEEE Journal of Selected Topics in Signal Processing*, **12**, 1144-1154. <https://doi.org/10.1109/JSTSP.2018.2872460>
- [35] Boyd, S., Parikh, N., Chu, E., Peleato, B. and Eckstein, J. (2011) Distributed Optimization and Statistical Learning via the Alternating Direction Method of Multipliers. *Foundations and Trends® in Machine Learning*, **3**, 1-122. <https://doi.org/10.1561/22000000016>
- [36] Tang, W., Hu, J., Zhang, H., *et al.* (2015) Kappa Coefficient: A Popular Measure of Rater Agreement. *Shanghai Archives of Psychiatry*, **27**, 62-67.
- [37] 李鸿吉. 模糊数学基础及实用算法[M]. 北京: 科学出版社, 2005.