

基于深度度量学习的导弹气动系数预测

刘 林¹, 杨春明¹, 蔺佳哲², 向宏辉³

(1. 西南科技大学计算机科学与技术学院, 绵阳 621000; 2. 中国空气动力研究与发展中心计算空气动力研究所, 绵阳 621000; 3. 中国航发四川燃气涡轮研究院, 绵阳 621000)

摘要: 传统多输出深度神经网络在导弹气动性能系数预测任务中, 通常采用均方误差 (Mean square error, MSE) 和平均绝对误差 (Mean absolute error, MAE) 来训练网络, 但在小样本及无物理方程约束的情况下, MSE 与 MAE 对导弹性能系数之间的约束和不同导弹样本之间的区分就会降低。针对该问题, 提出一种基于深度度量学习的 K 最近邻大边距损失函数 (K-nearest neighbor large margin, KNNLM), 它通过边距约束将大差异输出样本推开, 拉近相近输出样本, 以此来解决样本及样本间的约束区分问题。以导弹气动外形及工况参数作为输入, 4 种气动系数作为输出, 在反向传播神经网络 (Backpropagation neural network, BPNN) 和多任务学习神经网络 (Multi-task learning neural network, MTLNN) 中分别采用 MSE、MAE、KNNLM 进行实验对比, 实验结果表明: KNNLM 在 BPNN 和 MTLNN 中的精度相比于 MSE 和 MAE 最大能够提升 14.44% 和 16.35%, 最少提升 3.72%。KNNLM 能够在少样本及无物理知识约束的情况下, 能更好地对导弹样本进行约束区分, 使深度神经网络模型的预测精度更高, 且鲁棒性更强。

关键词: 深度度量学习; 导弹; 气动性能预测; K 最近邻大边距; 多输出

中图分类号: V211.24

文献标志码: A

文章编号: 1005-2615(2024)05-0950-10

Missile Aerodynamic Coefficient Prediction Through Deep Metric Learning

LIU Lin¹, YANG Chunming¹, LIN Jiazhe², XIANG Honghui³

(1. School of Computer Science and Technology, Southwest University of Science and Technology, Mianyang 621000, China; 2. Computational Aerodynamic Research Institute, China Aerodynamics Research and Development Center, Mianyang 621000, China; 3. AECC Sichuan Gas Turbine Establishment, Mianyang 621000, China)

Abstract: Mean square error (MSE) and mean absolute error (MAE) are usually used to train traditional multi-output deep neural networks in missile aerodynamic coefficient prediction. However, in the case of small sample size and no physical equation constraint, the constraint between MSE and MAE on missile performance coefficient and the distinction between different missile samples will be reduced. A K nearest neighbor large margin (KNNLM) loss function based on deep metric learning is proposed. The method uses the margin constraint to push the output samples with large differences away, and close the similar output samples. Taking the aerodynamic shape and working condition parameters of the missile as input and four aerodynamic coefficients as output, MSE, MAE and KNNLM are used for experimental comparison in backpropagation neural network (BPNN) and multi-task neural network (MTLNN). The experimental results show that compared with MSE and MAE, KNNLM can improve the accuracy by 14.44% and 16.35% at most, and 3.72% at least in BPNN and MTLNN. And the KNNLM can better distinguish the missile samples in the case of fewer samples and no physical knowledge constraint, so that the prediction accuracy of the deep neural network model is higher and the robustness is stronger.

基金项目: 四川省科技厅重点研发项目(2021YFG0031); 先进航空动力创新工作站项目(HK CX2022-01-022)。

收稿日期: 2023-05-08; **修订日期:** 2023-10-08

通信作者: 杨春明, 男, 副教授, E-mail: yangchunming@swust.edu.cn。

引用格式: 刘林, 杨春明, 蔺佳哲, 等. 基于深度度量学习的导弹气动系数预测[J]. 南京航空航天大学学报, 2024, 56(5): 950-959. LIU Lin, YANG Chunming, LIN Jiazhe, et al. Missile aerodynamic coefficient prediction through deep metric learning[J]. Journal of Nanjing University of Aeronautics & Astronautics, 2024, 56(5): 950-959.

Key words: deep metric learning; missile; aerodynamic performance prediction; K-nearest neighbor large margin (KNNLM); multi-output

气动性能参数是评价飞行器在空气中运动性能的重要指标,传统的气动性能参数评估方法主要是风洞试验和飞行试验以及计算流体力学(Computational fluid dynamics, CFD)模拟。风洞试验和飞行试验虽然能得到真实的气动性能参数,但是需要较长的试验准备时间和高昂的成本,而CFD模拟虽然在流场预测、气动建模、控制优化和设计优化等方面取得了优异的成果^[1-5],但是对于气动性能参数预测也还是需要大量的计算资源与时间。目前基于数据驱动的方法通过机器学习或深度学习从大量历史数据中开展性能预测,能够改进上述传统方法耗时长与计算资源开销大等问题,加快导弹的外形设计速度和缩短研发周期。

导弹的几何外形和工况条件决定了导弹的气动性能参数(如轴向力系数、法向力系数和压心位置等)。因此,导弹气动性能参数预测是一个非线性的多输入多输出回归问题。经典多任务机器学习方法需要对每个输出维度的气动参数分别建模,如支持向量回归(Support vector regression, SVR)^[6-7]、高斯过程回归(Gaussian process regression, GPR)^[8]、Kriging回归^[9]等,不仅训练成本高,且未考虑样本多输出维度之间的相关性(气动性能之间的相互影响,如轴向力系数、法向力系数和压心位置满足的物理公式)。虽然每个输出维度的预测结果都很优异,但模型整体鲁棒性较弱,在预测时输出之间缺乏一定的逻辑关系。神经网络有着天然的单输出与多输出特性,在多输出任务中,能使模型多输出目标之间参数信息共享,从而找到多输出目标之间的相关性,以此提高模型整体鲁棒性^[10-11]。

在缺乏物理方程约束的情况下,使用深度神经网络拟合的模型能够使多维输出参数之间符合一定的逻辑规律。但指导深度神经网络进行回归任务的损失函数,如均方误差(Mean square error, MSE)与平均绝对误差(Mean absolute error, MAE)等,只是对单条样本进行拟合,而没有考虑邻近样本的影响,不能很好地对不同导弹样本进行约束区分,导致模型在少样本及无物理方程约束的情况下降低预测精度。深度度量学习是数据间相似性距离度量及类内约束,能够度量不同样本的相似度,考虑不同样本数据之间的相互影响,约束区分不同样本^[12]。基于此,提出K最近邻大边距(K-nearest neighbor large margin, KNNLM)损失函数来指导深度神经网络的训练,KNNLM能够在气

动系数预测回归任务中使模型的训练空间中具有相近输出的样本距离更近,大差异输出的样本距离更远,以此实现不同样本区分及对应样本约束,使模型鲁棒性更强,本文在特定导弹气动数据集上采用不同网络、不同样本数进行了验证实验,其结果表明KNNLM的预测精度优于MSE和MAE。

1 相关工作

在导弹的气动性能预测方法中,传统的风洞试验、飞行试验和数值计算随着飞行器研发速度加快,在研发初期,逐渐不能满足现如今的研发需求。基于数据驱动的机器学习模型不需要复杂的计算过程与庞大的计算资源开销就能够在导弹研发初期对导弹的气动性能参数做出初步估计,为后续研发提供参考。根据气动性能参数预测回归任务的多输入、多输出特性,目前存在的方法主要有经典机器学习和深度神经网络模型。

经典机器学习方法主要有支持向量回归、高斯过程回归和Kriging模型。其支持向量回归是最大化间隔的同时最小化损失,而对于非线性问题能够使用核函数来解决。Andrés-pérez^[13]采用线性回归与支持向量回归对翼型气动性能进行了预测研究,实验表明支持向量回归预测精度更高,另外文中对翼型数据进行分析,表明其气动性能数据的非线性特性。高斯过程回归是一种高斯过程先验的非参数化模型^[14]。Barnhart等^[6]和胡伟杰等^[8]将高斯过程回归应用在气动系数预测任务中且与支持向量回归、线性回归、逐步线性回归和随机森林等经典机器学习方法做了对比,证明了其高斯过程回归在预测中的优越性。夏露等^[9]和韩少强等^[15]分别提出基于自适应的Kriging模型和基于梯度增强Kriging模型,这两个模型在气动性能优化设计中都取得了良好的优化效果。从上述的研究中可知在气动性能预测与优化设计中,高斯过程回归与Kriging模型优于支持向量回归等经典回归模型。但传统高斯过程回归的时间复杂度为 $O(n^3)$,训练时间会随着训练样本数 n 增加呈现几何倍的增长,而本质上Kriging模型与高斯过程回归相似^[16]。另外在气动性能预测任务中,高斯过程回归与Kriging模型很难进行精确的多输出任务建模,其都是对每个输出目标,如升力系数、阻力系数分别建模预测。这样大大增加了训练时间,且没有考虑多输出目标之间和相似样本间的相互影响,导致模型鲁棒性不强。

在气动性能预测任务中,神经网络模型具有与高斯过程回归相似的预测精度,另外神经网络模型更是天然具有多输入多输出特性,以至于在气动性能预测任务中具有更大的优势,原智杰等^[17]和韩建福等^[18]在气动性能预测研究中,分别将遗传算法融入反向传播神经网络(Backpropagation neural network, BPNN)来避免人工参数选取的不确定性和采用双BPNN网络来弥补单一网络的不足。上述神经网络模型研究主要从模型参数角度考虑,没有从具体数据特性考虑。气动性能数据的输入是由几何参数与工况参数(马赫数、迎角等)组成,其具有不同的度量特性,输出之间也具备一定的物理知识关联。张骏等^[11]将翼型几何参数和工况参数看作两种大差异输入数据,将两种数据分别输入卷积神经网络(Convolutional neural network, CNN)和BPNN网络,形成多任务学习网络模型(Multi-task learning neural network, MTLNN)进行翼型气动系数预测。对于已知物理知识的运用,采用嵌入物理知识神经网络模型(Physics informed neural network, PINN)^[19-21],蔺佳哲等^[10]将MTLNN与PINN结合,将每个输出维度考虑成一个任务,再用PINN结合每个任务的输出计算物理知识损失,使得输出进一步符合物理知识的逻辑关系,其预测精度比传统MTLNN更高,模型鲁棒性也更强。而从模型的损失函数来看,这些深度神经网络模型都是采用MSE或者MAE,如果数据中所知的物理知识约束计算过于复杂^[19-21],PINN就很难嵌入模型中,而单独的MSE、MAE在计算损失时,对气动性能系数之间的约束和不同样本之间的区分就会降低。针对这一问题,其深度度量学习的损失函数能够在训练时考虑临近样本的影响程度,将大差异样本推远,拉近相似样本来进行同条样本约束和不同样本区分,最终增强模型的鲁棒性。本文使用的主要数学符号参考表1。

表1 主要变量及符号描述

Table 1 Description of main variables and symbols

符号	变量
D	训练数据集
θ	网络参数
k	最近邻数
$x_i \in \mathbf{R}^{1 \times d}$	特征样本
$y_i \in \mathbf{R}^{1 \times q}$	标签样本
$\ \cdot\ _p$	p范数
$(e(x_i; \theta), y_i, y_i)$	回归任务三元组
y	x_i 的K最近邻样本输出值集合
ξ_i	KNNLM中第二项的值,对应于 x_i

2 K最近邻大边距损失

2.1 问题定义

神经网络传统回归损失函数:假设训练数据 $D=(X, Y)=\{x_i, y_i\}_{i=1}^n$, n 表示训练数据条数,每条数据 $x_i \in \mathbf{R}^{1 \times d}$ 为 d 维的输入变量和 $y_i \in \mathbf{R}^{1 \times q}$ 为 q 维的输出变量, d 和 q 分别代表输入特征数和输出变量数。 X 、 Y 分别为大小 $n \times d$ 和 $n \times q$ 的矩阵。在气动系数预测任务中,其目标是将训练样本特征 x_i 输入神经网络,形成 q 维的嵌入映射 $f(x_i; \theta) \rightarrow y_i$ 来预测输出目标的值, θ 表示神经网络学习的嵌入参数,下文将以 $e(x_i; \theta)$ 来表示 $f(x_i; \theta)$,即

$$\mathcal{L} = \frac{1}{n} \sum_{i=1}^n \|e(x_i; \theta) - y_i\|_p^2 \quad (1)$$

式中: $\|\cdot\|_p$ 表示p范数,在回归预测中所期望的是使式(1)中损失最小 $\arg \min \|e(x_i; \theta) - y_i\|_p^2$,优化目标函数使用经典的梯度下降算法。具体的MSE和MAE如下

$$\text{MSE} = \frac{1}{n} \sum_{i=1}^n \|e(x_i; \theta) - y_i\|_2^2 \quad (2)$$

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^n \|e(x_i; \theta) - y_i\|_1 \quad (3)$$

MSE和MAE指导神经网络执行端到端的气动系数预测回归任务时,其能快速直接地拟合对应输出目标,但在拟合过程中,根据损失函数对模型的制约,模型只执行了当前样本特征 x_i 与 y_i 的拟合。当样本数据过少,损失函数对不同样本的约束与区分就会降低。因此存在引入物理知识约束形成PINN模型,提高导弹输出系数之间的约束,增加预测精度。但真实气动数据集是稀少且珍贵的,另外对于每个数据集,不一定都能找到合适的物理知识进行嵌入。在这种少样本和无物理知识约束的情况下,要使训练的深度神经网络模型进行更好的气动数据样本约束和区分,增加模型的鲁棒性,这样的要求对当前模型来说是困难的。

在用于气动系数预测任务且当样本数据过少和无物理知识约束时,模型对于多输出目标样本之间缺乏约束区分,从而导致训练的模型鲁棒性降低。根据目前损失函数存在的缺陷,引入深度度量学习,其损失函数能考虑不同样本之间的相似性,进行样本距离度量,以此来约束区分不同样本。

传统损失函数指导深度神经网络进行分类任务学习,不同类具有较近相关性时,其网络学习到的嵌入空间有较大可能将不同类划分在相同空间中。采用深度度量学习,能够在正确分类的同时使

得网络空间将同类样本拉得更近,不同类样本推得更远,进行更好的类内约束和类间区分。如在图像分类领域^[22-23],深度度量学习展现出了显著优势,其通过深度神经网络(如CNN)进行特征嵌入,然后根据嵌入提取到的特征在某个维度进行图片之间的相似度比较,从而进行分类,而这个维度就是需要深度度量学习的维度。在深度度量学习中,经典的损失函数有:三元组损失^[24]、代理损失^[25]等。

三元组损失函数都具有一个锚点 $x_i^{(a)}$ (Anchor)、一个正样本点 $x_j^{(p)}$ (Positive) 和一个负样本点 $x_z^{(n)}$ (Negative), 其中正样本点与锚点属于同一类别, 负样本点与锚点属于不同类别, 其目标是在嵌入空间中使得不同类别样本点距离更远, 相同类别样本点距离更近, 具体如图1所示。

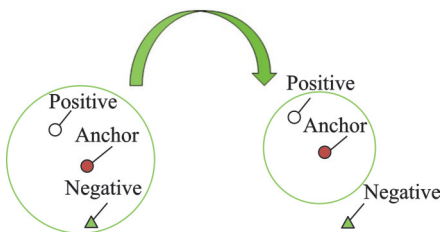


图1 深度度量学习三元组损失

Fig.1 Triplet loss of deep metric learning

其公式为

$$e(x_i^{(a)}; \theta) - e(x_j^{(p)}; \theta) < e(x_i^{(a)}; \theta) - e(x_z^{(n)}; \theta) \quad (4)$$

期望 $\arg \min(e(x_i^{(a)}; \theta) - e(x_j^{(p)}; \theta))$ 和 $\arg \max(e(x_i^{(a)}; \theta) - e(x_z^{(n)}; \theta))$ 。

为了增加函数的稳定性,使得相同类更近,不同类更远,通常会加入一个常数 m 及 $\text{margin}(m > 0)$ 来扩大式(4)中两边的差值。当锚点与正样本点的距离加上 m 后其值小于锚点与负样本的距离就认为其损失为0,相反,认为损失为两者差值,即

$$\sum \left[\|e(x_i^{(a)}; \theta) - e(x_j^{(p)}; \theta)\|_2^2 - \|e(x_i^{(a)}; \theta) - e(x_z^{(n)}; \theta)\|_2^2 + m \right]_+ \quad (5)$$

式中 $[\cdot]_+ = \max(\cdot, 0)$ 。从式(4,5)可以知道,深度度量学习的损失函数能够考虑当前样本及与其相似样本的关联性。

2.2 模型定义

深度度量学习的损失函数中,通常需要一个三元组,如式(5)所示。而气动系数预测任务本质属于一种回归任务,数据中不存在类别之分(也可以说每条样本属于一种类别,不存在相同类别)。受Liu等^[26]研究启发,本文将锚点对应的正样本点归为同一条样本,即锚点就是正样本点,而其余样本为负样本点。另外回归任务是对真实数据的拟合,

所以网络嵌入映射主要是拟合目标输出值,如式(1)所示,为了使网络的映射嵌入接近当前样本对应的输出,远离其余样本对应的输出,即

$$\|e(x_i; \theta) - y_i\|_2^2 < \|e(x_i; \theta) - y_j\|_2^2 \quad (6)$$

式中 y_j 为属于除 y_i 在内的其余样本对应的真实输出目标值,所以新的三元组为 $(e(x_i; \theta), y_i, y_j)$, 本文期望 $\arg \min(e(x_i; \theta) - y_i)$ 和 $\arg \max(e(x_i; \theta) - y_j)$ 。尽可能使嵌入映射拟合当前对应样本值的同时拉开不对应样本。同式(5),为了增加式(6)中两边的差值,加入一个边距变量。但是本文中的边距变量不是一个固定常数,而是随着不同样本的输出值变化而变化的。为了进一步增强网络区别不同样本的能力和保持其一定的包容性,本文选取 k 个最近邻样本。

$$\sum_{i=1}^n \left[\|e(x_i; \theta) - y_i\|_2^2 - \|e(x_i^{(a)}; \theta) - y\|_2^2 + \Delta(y_i, y) \right]_+ \quad \forall y \in \text{Nei}(i), \forall i \quad (7)$$

式中 $\text{Nei}(i)$ 为样本 i 的最近邻样本。由于回归任务中,目标输出值之间是真实变动,不像分类任务一样有固定类别之分,所以式(5)中的 margin 采用 $\Delta(\cdot)$ 代替,表示两个目标之间的L2范数,根据两个样本输出值之间的远近来调整其惩罚力度大小,距离越远其惩罚值越大,相反惩罚值越小,使其损失更加灵活,更加适用于回归任务。具体如式(7)所示,使网络不仅仅考虑当前样本,在拟合过程中,还会根据样本及样本周围的数据进行综合考虑,增加相差过大输出值样本之间的距离,降低样本特征相似但输出相差过大的影响。

在气动系数预测任务中,当数据样本过少且无物理知识约束时,度量学习能够增强模型输出值之间的逻辑关系,对不同气动数据样本进行约束和区分,提高模型鲁棒性。为了加强式(7)中当前样本输入与输出之间的拟合关系,将式(7)与式(2)结合,最终KNNLM损失函数为

$$\begin{aligned} \mathcal{L} &= \frac{1}{n} \sum_{i=1}^n \|e(x_i; \theta) - y_i\|_2^2 + \frac{c}{n} \sum_{i=1}^n \xi_i \\ \text{s.t. } &\|e(x_i; \theta) - y_i\|_2^2 + \Delta(y_i, y) - \xi_i < \|e(x_i^{(a)}; \theta) - y\|_2^2 \\ &\forall y \in \text{Nei}(i), \forall i, \xi_i \geq 0 \end{aligned} \quad (8)$$

式中 ξ_i 为最近邻样本的差异值,为式(7)的计算值。

2.3 模型分析

在式(8)中,将式(2,7)组合从而分为了两个部分,其第1项为基础损失函数,第2项为度量区分函数。这样在基础损失函数后面加上最近邻对比情况,不仅考虑了当前样本对应输出的拟合,也

考虑了临近样本输出的影响。进一步,为了平衡前后两项的权重关系,使后一项乘以常数 C ,其 C 值根据实际情况调整,决定了第2项中模型对最近邻样本的关注度。训练过程如算法1所示。

算法1 K最近邻损失训练

输入: D :训练数据 $D=(X,Y)=\{x_i,y_i\}_{i=1}^n$,初始化参数 θ,k,c

for $i=1$ to n do

基于欧氏距离获得 x_i 的 k 个最近邻样本。形成回归任务三元组 $(e(x_i;\theta),y_i,y)$,计算三元组损失。

$$\xi_i \leftarrow \max\{0, \max(e(x_i;\theta), y_i, y)\}$$

end for

$$\mathcal{L} = \frac{1}{n} \sum_{i=1}^n \|e(x_i;\theta) - y_i\|_2^2 + \frac{c}{n} \sum_{i=1}^n \xi_i$$

return \mathcal{L}

3 实验与结果分析

3.1 实验数据及比较方法

本文实验数据来源文献[10],由于本文主要探究当前导弹气动系数预测研究,数据样本过少且无先验知识约束的情况,所以只选取了4种不同外形参数的导弹数据,总计2 880条。数据中包含13种几何外形特征参数,3种工况参数(马赫数 Ma 、迎角 α 、滚转角 φ),4种输出气动特性参数(轴向力系数 C_A 、法向力系数 C_N 、俯仰力矩系数 C_m 、压心位置 X_p),4种导弹主要外形差异部分如表2所示。表中:LF为后翼距弹头距离;CRF为后翼根弦长度;CTF为梢弦长度;BKF为后掠长度;BF为半展长。

表2 4种导弹几何外形参数差异部分

Table 2 Differences in geometric parameters of four types of missiles

序号	LF	CRF	CTF	BKF	BF
1	1.593 6	0.056 4	0.027 6	0.028 8	0.084
2	1.608	0.042	0.042	0	0.084
3	1.53	0.12	0	0.12	0.06
4	1.566	0.084	0.036	0.048	0.06

在训练过程中,为了探究少样本情况,随机选取500、1 500和2 500个样本作为不同样本量数据,以此探究提出的KNNLM损失函数相比于MSE、MAE在少样本和无物理知识约束情况下的优异性。为了弥补数据随机选取带来的差异和尽可能保证模型稳定性的情况下,在每种数据量的实验中进行10次实验取平均值作为最终参考评价。另外每组数据的训练集、验证集、测试集均随机划分为8:1:1,神经网络训练的mini_batch大小为30,且网络输出维度均为4。

在深度学习方面,主要采用BPNN和MTLNN进行实验探究,训练神经网络的损失函数采用传统的MAE、MSE和本文提出的KNNLM,回归评价指标采用均方根误差 $RMSE = \frac{1}{n} \sum_{i=1}^n \sqrt{(e(x_i;\theta) - y_i)^2}$ 。

具体网络情况介绍如下:

(1)MTLNN网络使用文献[10]中的网络结构,但是不考虑物理知识约束的情况,不采用内嵌物理知识PINN。

(2)BPNN根据数据的不同具有不同网络层数和节点数,因此将在5组实验中探究采用MSE和MAE损失函数的最优网络结构,然后将其最优RMSE结果的网络结构用于KNNLM,以此实现相同网络结构和不同损失函数的公平对比结果,证明提出的损失函数的有效性。

实验采用Pytorch深度学习框架,在计算机配置为Intel(R)Core(TM)i7-8700 CPU@3.20 GHz,3.19 GHz上进行实验。

3.2 BPNN结构探索

对于反向传播神经网络的网络结构探索,为了避免网络过拟合和欠拟合,首先从3层网络架构、节点数为(32,16,8)开始进行实验,然后依次递增网络层数及每层节点数,损失函数分别选取MSE和MAE。具体实验在MSE损失函数和MAE损失函数且样本数为500、1 500、2 500的情况下进行,如表3所示。

图2和图3是根据表3所得到的MSE和MAE损失函数下的BPNN不同网络结构、不同样本数的RMSE值折线图。其中图2横坐标表示从3层网络(32,16,8)开始依次递增的网络层数(如:6表示6层网络(256,128,64,32,16,8)),图3横坐标表示在当前层数时基于原始节点数的 2^{n-1} 倍(如:4_1为4层网络结构(64,32,16,8),5_1为5层网络结构(128,64,32,16,8),4_3就是在4_1网络节点数上的 2^2 倍即(256,128,64,32))。

由表3、图2和图3可知:(1)当损失函数为MSE且样本数为500和2 500时,最佳的网络结构分别为(128,64,32,16)、(256,128,64,32)的4层结构。在样本数为1 500时,最佳网络结构为(256,128,64,32,16)的5层网络结构。在500、1 500和2 500个样本数时各自的最优预测RMSE值为0.074 52、0.038 16、0.026 59。(2)当损失函数为MAE时,最佳网络结构层数都为4层网络结构,500、1 500个样本数时最佳节点数都为(128,64,32,16),而样本数为2 500时,最佳的网络节点数为(256,128,64,32)。500、1 500、2 500个样本数

在 MAE 损失函数和各自的网络结构下的最优

随着样本数量的增加,不管是MSE损失函数还是

RMSE 值分别为 0.072 41、0.037 08 和 0.029 69。

MAE 损失函数,预测的最优RMSE值都在减少。

表 3 BPNN 网络结构探索

Table 3 Network structure exploration of BPNN

损失函数	样本数	网络结构	RMSE	损失函数	样本数	网络结构	RMSE
MSE	500	(32,16,8,4)	0.096 15	MAE	500	(32,16,8,4)	0.099 87
		(64,32,16,8,4)	0.089 39			(64,32,16,8,4)	0.072 65
		(128,64,32,16,8,4)	0.090 27			(128,64,32,16,8,4)	0.078 68
		(128,64,32,16,4)	0.074 52			(128,64,32,16,4)	0.072 41
		(256,128,64,32,4)	0.084 28			(256,128,64,32,4)	0.082 41
	1 500	(32,16,8,4)	0.049 42		1 500	(32,16,8,4)	0.053 81
		(64,32,16,8,4)	0.042 96			(64,32,16,8,4)	0.039 87
		(128,64,32,16,8,4)	0.040 55			(128,64,32,16,8,4)	0.054 20
		(256,128,64,32,16,8,4)	0.059 53			(128,64,32,16,4)	0.037 08
		(256,128,64,32,16,4)	0.038 16			(256,128,64,32,4)	0.039 65
	2 500	(512,256,128,64,32,4)	0.043 89		2 500	(32,16,8,4)	0.047 51
		(32,16,8,4)	0.043 38			(64,32,16,8,4)	0.039 50
		(64,32,16,8,4)	0.034 58			(128,64,32,16,8,4)	0.039 75
		(128,64,32,16,8,4)	0.035 42			(128,64,32,16,4)	0.034 47
		(128,64,32,16,4)	0.030 33			(256,128,64,32,4)	0.029 69
		(256,128,64,32,4)	0.026 59			(512,256,128,64,4)	0.030 84
		(512,256,128,64,4)	0.026 91				

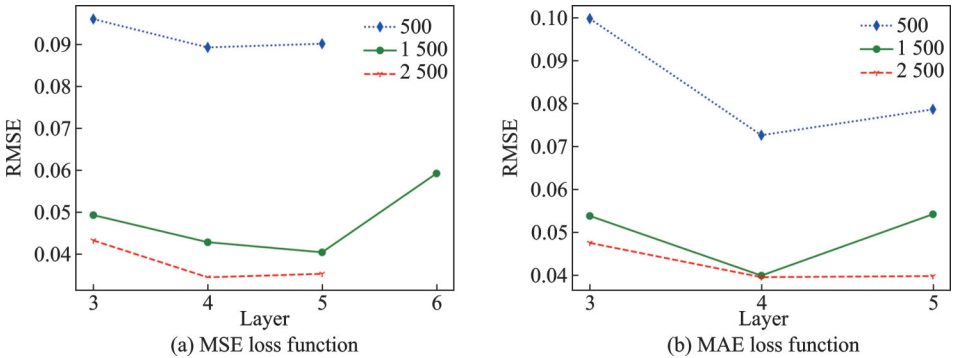


图 2 BPNN 网络层数

Fig.2 BPNN layers

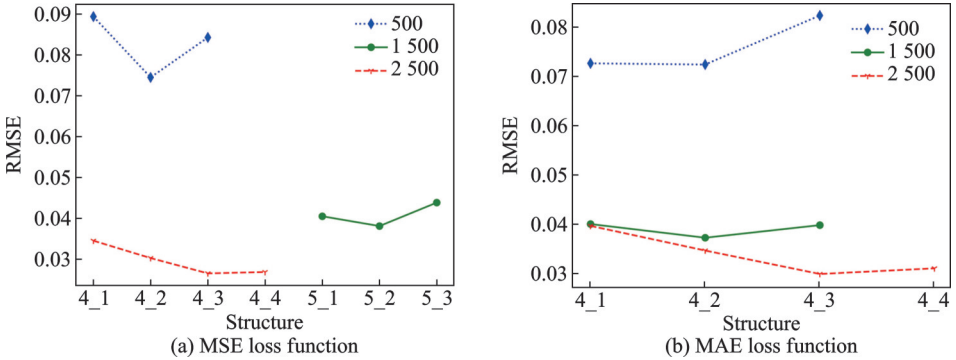


图 3 BPNN 网络节点数

Fig.3 BPNN structure

在 1 500 个样本数时,MSE、MAE 损失函数对

应的最优网络层数不同,可能原因是 MAE 在少样

本情况下要优于 MSE 损失函数,因为在 500、1 500

个样本数时最优的 RMSE 值对应的都是 MSE 损

失函数。MSE 进行平方而 MAE 取绝对值,在预测

系数中最小值为 0.03,最大值达到了 10 左右,MSE

损失函数的平方会拉大损失值,所以 MSE 损失函

数在 1 500 个样本数时需要的网络层数更多。当样

本数在500、2 500时,对于模型来说,样本数较少或者足够大,其MSE、MAE损失函数之间差距不明显。

3.3 KNNLM 参数分析

3.3.1 KNNLM的C值分析

在KNNLM函数中,第2项表示临近样本对的影响程度,而其中的导弹气动系数预测为回归任务,KNNLM第1项占比应该大于第2项,多个临近样本对的影响权重不应该超过原始项,所以在分析中可知KNNLM第2项权重 C 占比应该维持在0.1左右。进一步对其做分析实验,取 $C=0.001$ 、0.01、0.1、1、10和100。又因为实验采用的mini_batch大小为30,所以固定 $K=8$ 。图4中每个子图的纵坐标表示精度RMSE值,横坐标表示权重参数 C 值变化,其中每条不同颜色的折线表示在BPNN网络和MTLNN网络下采用KNNLM训练时的不同样本数。图4(a)中的1 500_1和1 500_2分别表示1 500个样本数时所对应的网络结构(128,64,32,16)和(256,128,64,32,16)。

从图4可知在两个网络中,1 500和2 500个样本数时,随着平衡权值参数 C 值的增加,RMSE值总体呈现上升趋势,网络预测精度逐渐下降。在500个样本数时,两个网络的RMSE值曲线随着 C 值的增加先下降然后上升,可看出在 $C=0.1$ 附近的RMSE值最优。另外在 $C=0.1$ 时,随着样本数的增加,两个网络附近的RMSE值变化也在趋于

平缓。MTLNN网络呈现出 C 值越小、RMSE值越低的趋势,BPNN在 C 值小于0.1时,RMSE值呈现上升趋势。总体来说,不管是MTLNN网络还是BPNN网络,在3种样本数下, C 值在0.001~0.1之间其RMSE值最优,进一步证实了分析中KNNLM第2项权重不应大于第1项的正确性。

3.3.2 KNNLM的K值分析

由式(8)可知, y 为 x_i 的 K 最近邻样本的目标值及KNNLM使用 K 来确定实例邻居个数,所以为了探究 K 值的具体影响程度,再由3.3.1节对 C 值探究结果可知,取 $C=0.1$, $K=6,8,10,12,14$ 。在图5中,使用不同数量的3组样本来探究 K 值变化对KNNLM损失函数在不同网络中的影响程度。图5(a)例中1 500_1和1 500_2分别表示BPNN的网络结构为(256,128,64,32,16)和(128,64,32,16)。

从图5可知,在500个样本数时两个网络在 $K=10$ 时最优。在1 500个样本数时,BPNN网络中 $K=8$ 时,RMSE最小,而MTLNN中 $K=10$,RMSE最小。在2 500个样本数时 K 的值在8~10之间,RMSE上升之后又下降,但变化幅度较小。总体来说,BPNN网络中RMSE随着 K 值的增加呈现上升趋势,而MTLNN网络中RMSE随着 K 值变化而呈现出波动变化。综上所述,在观察了整体变化趋势后,KNNLM在 $C=0.1$ 、 $K=8$ 或10是较理想的一种参数取值。

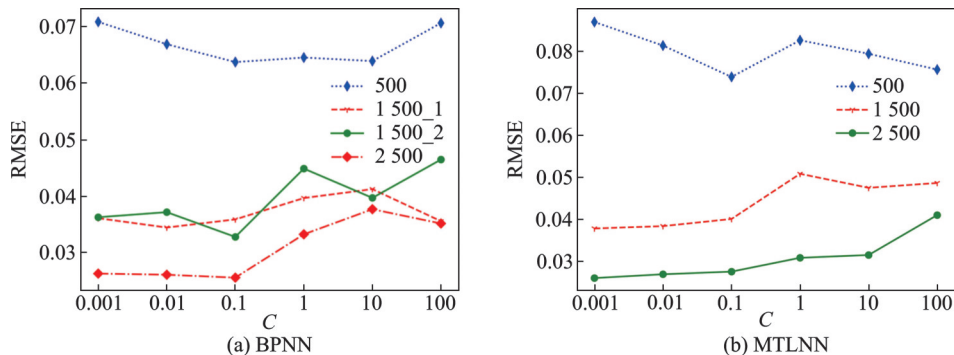


图4 KNNLM的C值变化

Fig.4 Change of C value of KNNLM

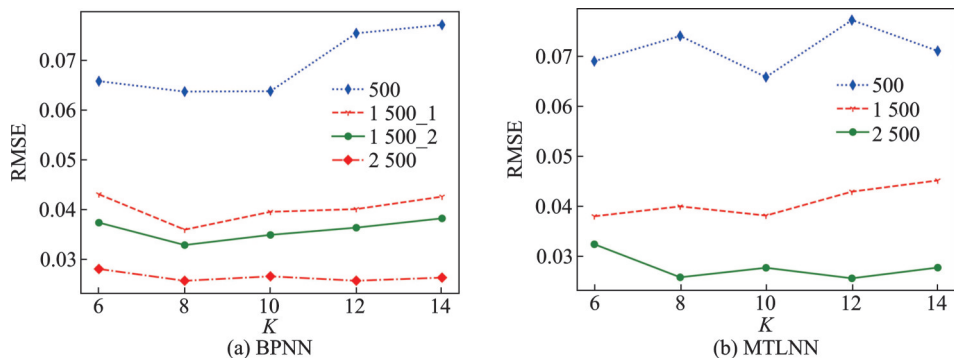


图5 KNNLM的K值变化

Fig.5 Change of K value of KNNLM

3.4 实验结果分析

由 3.3 节分析可知, $C=0.1$, K 取 8~10 对于 BPNN 和 MTLNN 网络最优。表 4 展示了不同样本数时所对应的不同损失在两种网络中的实验结果。每一行表示不同样本数, 每一列表示在不同样本数下, 不同损失函数和网络。在所有方法中, 对于不同网络结构下, 采用不同的损失时, 最好的结果以粗体表示, 基于这些实验结果可以得到如下结论:

(1) 从表 4 可看出, 在两种网络中, KNNLM 都要优于 MSE 和 MAE 损失函数。在 BPNN 网络中, 当样本数为 500 和 1 500 时, MAE 损失函数优于 MSE 损失函数, 而在 2 500 个样本数时, MSE 优于 MAE。对于 MTLNN 网络, 在 3 组样本数下,

MSE 损失函数都要优于 MAE 损失函数。

(2) 表 5 为 KNNLM 对比 MSE 和 MAE 的精度提升百分比。可知, KNNLM 与 MSE 对比, 其在 BPNN 和 MTLNN 中都是随着样本数的增加其精度百分比呈现递减趋势。KNNLM 与 MAE 的对比情况中, 除了在 2 500 个样本数时的提升精度较高外, 500~1 500 个样本数其提升精度都呈现下降趋势。从总体上看, 在 BPNN 网络和 MTLNN 网络中, KNNLM 的最大精度提升百分比是在 500 个样本数时, 其分别为 14.44%、16.35%。综上所述, KNNLM 能够在少样本且无物理知识约束的情况下, 能更好地进行导弹气动系数约束和不同导弹样本区分, 最大限度地提升模型的预测精度和鲁棒性。

表 4 不同方法的 RMSE 值实验结果

Table 4 Experimental results of RMSE value of different methods

样本数	MSE_BPNN	MAE_BPNN	KNNLM_BPNN	MSE_MTLNN	MAE_MTLNN	KNNLM_MTLNN
500	0.074 52	0.072 41	0.063 76	0.077 24	0.078 51	0.065 67
1 500	0.038 16	0.037 08	0.032 81	0.043 21	0.044 89	0.037 78
2 500	0.026 59	0.029 69	0.025 60	0.028 93	0.030 69	0.025 79

表 5 KNNLM 对比 MSE 和 MAE 的精度提升百分比

Table 5 Accuracy improvement percentage of KNNLM compared with MSE and MAE %

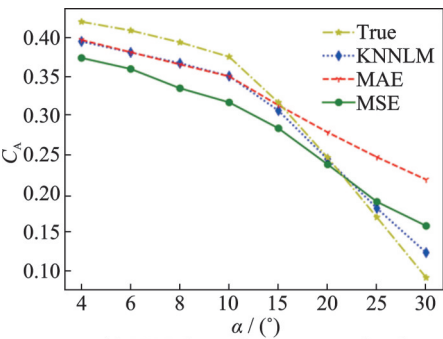
样本数	BPNN		MTLNN	
	KNNLM_MSE	KNNLM_MAE	KNNLM_MSE	KNNLM_MAE
500	14.44	11.95	14.9	16.35
1 500	14.02	11.52	12.57	15.84
2 500	3.72	13.78	10.85	15.97

3.5 KNNLM 预测应用效果对比

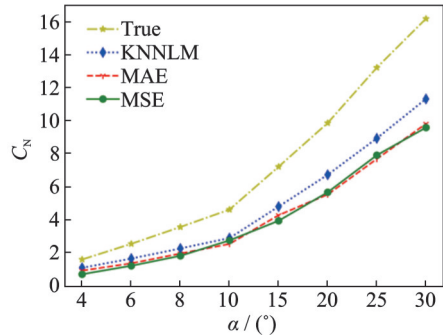
为了进一步展示 KNNLM 指导深度神经网络在少样本且无物理约束情况下的预测效果, 且由于样本数 1 500 与 2 500 的整体 RMSE 值差距小于样本数 500 与 1 500 的 RMSE 值差距, 为了进行更加明显的小样本预测区分, 所以取 500 个样本数训练的模型进行预测。另外由实验结果可知, 在少样本且无物理约束情况下, 无论是 MAE、MSE 还是 KNNLM 损失函数, BPNN 网络都优于 MTLNN 网络, 所以采用 BPNN 网络来进一步说明 KNNLM、MAE、MSE 的对比效果, 随机选取了 8 组训练数据之外的导弹气动数据进行实验展示。

如图 6 所示, 分别是 4 种气动系数 (C_A 、 C_N 、 C_{mz} 、 X_p) 的预测曲线图, 可知, KNNLM 对 C_A 、 C_N 、 C_{mz} 的预测值更接近真实值, 预测效果优于 MSE、MAE。对于 X_p 系数, KNNLM 与 MSE 的预测效果相距不大, MAE 更优。对于 4 种气动系数的预测效果, KNNLM 在整体上要优于 MAE、MSE, 这说

明在少样本且无物理约束的情况下, KNNLM 能够更适合多输出的导弹气动系数预测任务。



(a) BPNN's prediction curves for C_A



(b) BPNN's prediction curves for C_N

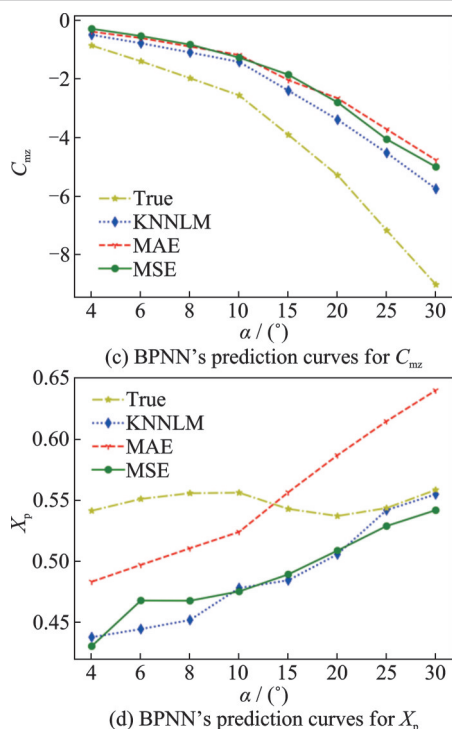


图6 BPNN网络气动系数预测曲线

Fig.6 Aerodynamic coefficient prediction curves of BPNN

4 结 论

针对传统损失函数MSE和MAE指导深度神经网络进行导弹气动系数预测任务存在的局限,基于深度度量学习提出一种K最近邻大边距损失函数(KNNLM)来指导深度神经网络的训练:

(1) 探究了KNNLM与MSE、MAE在BPNN网络与MTLNN网络中且在3组不同样本数的训练数据下的预测精度;其结果表明,KNNLM在多输出BPNN中比MSE和MAE在预测精度上能够提升3.72%~14.44%。在MTLNN网络中比MSE和MAE在预测精度上能够提升10.85%~16.35%。

(2) 在500个样本数时,KNNLM在两种网络上带来的预测精度提升百分比最大,其分别为14.44%和16.35%。这表明KNNLM在少样本及无物理知识约束的情况下,能够综合考虑实例与临近样本之间的相似性,从而约束区分不同导弹样本,以此拉近相似输出样本而推开差距过大输出样本对,能够较大程度避免在相近输入时不同输出带来的影响,从而提高深度神经网络模型的预测精度和鲁棒性。

参考文献:

[1] WHITE C, USHIZIMA D, FARHAT C. Fast neural network predictions from constrained aerodynamics datasets[C]//Proceeding of AIAA Scitech 2020 Fo-

rum. Orlando, FL: American Institute of Aeronautics and Astronautics, 2020.

[2] RABAULT J, REN F, ZHANG W, et al. Deep reinforcement learning in fluid mechanics: A promising method for both active flow control and shape optimization[J]. Journal of Hydrodynamics, 2020, 32(2): 234-246.

[3] 张伟伟, 寇家庆, 刘溢浪. 智能赋能流体力学展望[J]. 航空学报, 2021, 42(4): 26-71.

ZHANG Weiwei, KOU Jiaqing, LIU Yilang. Prospect of artificial intelligence empowered fluid mechanics[J]. Acta Aeronautica et Astronautica Sinica, 2021, 42(4): 26-71.

[4] BRUNTON S L, NOACK B R, KOUMOUTSAKOS P. Machine learning for fluid mechanics[J]. Annual Review of Fluid Mechanics, 2020, 52(1): 477-508.

[5] BAIGES J, CODINA R, CASTAÑAR I, et al. A finite element reduced-order model based on adaptive mesh refinement and artificial neural networks[J]. International Journal for Numerical Methods in Engineering, 2020, 121(4): 588-601.

[6] BARNHART S A, NARAYANAN B, GUNASEKARAN S. Blown wing aerodynamic coefficient predictions using traditional machine learning and data science approaches[C]//Proceeding of AIAA Scitech 2021 Forum. Virtual Event: American Institute of Aeronautics and Astronautics, 2021.

[7] 彭博, 聂蓉梅, 陈海东. 基于支持向量机的火箭气动力学代理模型构建方法[J]. 导弹与航天运载技术, 2013(4): 33-37.

PENG Bo, NIE Rongmei, CHEN Haidong. Surrogate model construction for rocket aerodynamic discipline based on vector machine[J]. Missiles and Space Vehicles, 2013(4): 33-37.

[8] 胡伟杰, 黄增辉, 刘学军, 等. 基于自动核构造高斯过程的导弹气动性能预测[J]. 航空学报, 2021, 42(4): 289-302.

HU Weijie, HUANG Zenghui, LIU Xuejun, et al. Missile aerodynamic performance prediction of gaussian process through automatic kernel construction[J]. Acta Aeronautica et Astronautica Sinica, 2021, 42(4): 289-302.

[9] 夏露, 王丹, 张阳, 等. 基于自适应代理模型的气动优化方法[J]. 空气动力学报, 2016, 34(4): 433-440.

XIA Lu, WANG Dan, ZHANG Yang, et al. Aerodynamic optimization method based on adaptive surrogate model[J]. Acta Aerodynamica Sinica, 2016, 34(4): 433-440.

[10] 蒯佳哲, 周岭, 武频, 等. 基于PIMTLNN的导弹气动特性快速预测技术研究[J]. 北京航空航天大学学

- 报, 2023, 49(10): 2669-2678.
- LIN Jiazhe, ZHOU Ling, WU Pin, et al. Research on rapid prediction technology of missile aerodynamic characteristics based on PIMTLNN[J]. Journal of Beijing University of Aeronautics and Astronautics, 2023, 49(10): 2669-2678.
- [11] 张骏, 张广博, 程艳青, 等. 一种气动大差异性数据多任务学习方法[J]. 空气动力学学报, 2021, 39(S): 1-10.
- ZHANG Jun, ZHANG Guangbo, CHENG Yanqing, et al. A multi-task learning method for large discrepant aerodynamic data[J]. Acta Aerodynamica Sinica, 2021, 39(S): 1-10.
- [12] 刘冰, 李瑞麟, 封举富. 深度度量学习综述[J]. 智能系统学报, 2019, 14(6): 1064-1072.
- LIU Bing, LI Ruilin, FENG Jufu. A brief introduction to deep metric learning[J]. CAAI Transactions on Intelligent Systems, 2019, 14(6): 1064-1072.
- [13] ANDRÉS-PÉREZ E. Data mining and machine learning techniques for aerodynamic databases: Introduction, methodology and potential benefits[J]. Energies, 2020, 13(21): 5807.
- [14] RASMUSSEN C E, WILLIAMS C K I. Gaussian processes for machine learning [M]. Cambridge, Mass: MIT Press, 2006.
- [15] 韩少强, 宋文萍, 韩忠华, 等. 基于梯度增强型Kriging模型的气动反设计方法[J]. 航空学报, 2017, 38(7): 138-152.
- HAN Shaoqiang, SONG Wenping, HAN Zhonghua, et al. Aerodynamic inverse design method based on gradient-enhanced Kriging model[J]. Acta Aeronautica et Astronautica Sinica, 2017, 38(7): 138-152.
- [16] 韩忠华. Kriging模型及代理优化算法研究进展[J]. 航空学报, 2016, 37(11): 3197-3225.
- HAN Zhonghua. Kriging surrogate model and its application to design optimization: A review of recent progress[J]. Acta Aeronautica et Astronautica Sinica, 2016, 37(11): 3197-3225.
- [17] 原智杰, 张公平, 崔茅, 等. 基于神经网络的导弹气动参数预测[J]. 航空兵器, 2020, 27(5): 28-32.
- YUAN Zhijie, ZHANG Gongping, CUI Mao, et al. Prediction of missile's aerodynamic parameters based on neural network[J]. Aero Weaponry, 2020, 27(5): 28-32.
- [18] 韩建福, 杜昌平, 叶志贤, 等. 基于双BP神经网络的扑翼飞行器气动参数辨识[J]. 计算机应用, 2019, 39(S2): 299-302.
- HAN Jianfu, DU Changping, YE Zhixian, et al. Identification of aerodynamic parameters of flapping-wing micro aerial vehicle based on double BP neural network [J]. Journal of Computer Applications, 2019, 39(S2): 299-302.
- [19] RAISSI M, PERDIKARIS P, KARNIADAKIS G E. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations[J]. Journal of Computational Physics, 2019, 378: 686-707.
- [20] KARNIADAKIS G E, KEVREKIDIS I G, LU L, et al. Physics-informed machine learning[J]. Nature Reviews Physics, 2021, 3(6): 422-440.
- [21] YUAN L, NI Y Q, DENG X Y, et al. A-PINN: Auxiliary physics informed neural networks for forward and inverse problems of nonlinear integro-differential equations[J]. Journal of Computational Physics, 2022, 462: 111260.
- [22] ZHENG W, WANG C, LU J, et al. Deep compositional metric learning[C]//Proceeding of 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Nashville, TN, USA: IEEE, 2021: 9316-9325.
- [23] ELEZI I, VASCON S, TORCINOVICH A, et al. The group loss for deep metric learning[C]//Proceeding of Computer Vision—ECCV 2020: 16th European Conference. Glasgow, UK: Springer International Publishing, 2020: 277-294.
- [24] SCHROFF F, KALENICHENKO D, PHILBIN J. FaceNet: A unified embedding for face recognition and clustering[C]//Proceeding of 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). [S.l.]: IEEE, 2015: 815-823.
- [25] MOVSHOVITZ-ATTIAS Y, TOSHEV A, LEUNG T K, et al. No fuss distance metric learning using proxies[C]//Proceedings of the IEEE International Conference on Computer Vision. [S.l.]: IEEE, 2017: 360-368.
- [26] LIU W, XU D, TSANG I W, et al. Metric learning for multi-output tasks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2019, 41(2): 408-422.

(编辑:刘彦东)