

基于双重模糊注意力机制的图像分类方法^{*}顾苏杭^{1,4} 王冶¹ 张远鹏³ 焦竹青^{1,2}(1. 常州大学王诤微电子学院 常州 213159; 2. 常州大学计算机与人工智能学院 常州 213159;
3. 南通大学医学院 南通 226001; 4. 常熟理工学院电气与自动化工程学院 苏州 215500)

摘要: 人类视觉系统在处理外界信息时,往往聚焦于目标的关键特征和结构,同时弱化非目标区域。此外,在经典的 CNN 模型中,图像中的噪声经逐层传播可能会干扰目标关键信息表征,导致无法准确提取特征。为此,本文提出一种基于双重模糊注意力机制的图像分类方法 DFAM-CNN。首先,针对 CNN 卷积层输出的特征图,通过引入模糊逻辑技术设计了模糊通道注意力机制和模糊空间注意力机制,并利用这两个机制在特征图的通道和空间方向上进行映射变换,生成与原特征图一一对应的重要模糊化特征图。其次,基于所有重要模糊化特征图,实现所有特征图通道权重和特征图内每个元素权重的计算,从而在通道和空间方向上突出与目标相关的特征。最后,通过模糊聚合操作对特征图进行降维,同时保留与目标相关的特征。为验证 DFAM-CNN 的有效性,在公开数据集 MedMNIST 和应用案例数据集上进行了大量的实验,实验结果验证了 DFAM-CNN 的有效性。特别地,与传统的最大池化方法相比,DFAM-CNN 在 BreastMNIST 和 DermaMNIST 子集上的准确率分别提升了 8.67% 和 7.40%。

关键词: 卷积神经网络;模糊逻辑技术;模糊通道注意力机制;模糊空间注意力机制

中图分类号: TP391.4; TN911.73 **文献标识码:** A **国家标准学科分类代码:** 520.2040

Image classification method based on dual fuzzy attention mechanism

Gu Suhang^{1,4} Wang Ye¹ Zhang Yuanpeng³ Jiao Zhuqing^{1,2}(1. Wang Zheng School of Microelectronics, Changzhou University, Changzhou 213159, China;
2. School of Computer and Artificial Intelligence, Changzhou University, Changzhou 213159, China;
3. Medical School, Nantong University, Nantong 226001, China;
4. School of Electrical and Automation Engineering, Changshu Institute of Technology, Suzhou 215500, China)

Abstract: The human visual system often focuses on the key features and structures of the target while weakening non-target areas when processing external information. In addition, in classical CNN models, noise in the image that propagates layer by layer may interfere with the representation of key information of the target, resulting in inaccurate feature extraction. Therefore, this article proposes an image classification method based on the dual fuzzy attention mechanism, named DFAM-CNN. Specifically, for the feature maps output by CNN convolutional layers, fuzzy channel attention mechanism and fuzzy spatial attention mechanism were first designed by introducing fuzzy logic technology. These two mechanisms were used to map and transform the feature maps along both the channel direction and spatial direction for generating important fuzzy feature maps that correspond to the original feature maps. Then, the channel weights of all feature maps and the weights of each element within each feature map were calculated based on all the determined important fuzzy feature maps, thereby highlighting the features related to the target in both the channel and spatial directions. Finally, dimensionality reduction was performed on the feature maps through fuzzy aggregation operations while retaining target-relevant features. To validate the effectiveness of DFAM-CNN, extensive experiments were conducted on both the public MedMNIST dataset and application-specific datasets. The experimental results validated the effectiveness of DFAM-CNN. Notably, compared with traditional max-pooling method, DFAM-CNN achieved accuracy improvements of 8.67% and 7.40% on the BreastMNIST and DermaMNIST subsets, respectively.

Keywords: convolutional neural network; fuzzy logic technology; fuzzy channel attention mechanism; fuzzy spatial attention mechanism

0 引言

图像分类一直是计算机视觉领域的研究热点之一,其

主要任务是根据图像的内容或特征,实现自动化分类并匹配预定义类别标签。随着机器学习和深度学习的发展,图像分类的识别性能取得了显著进步,其中,卷积神经网络

收稿日期:2025-01-08

* 基金项目:国家自然科学基金(62106025、51877013)、江苏省自然科学基金(BK20210940)、江苏省“333工程”(2024-3-0484)项目资助

(convolutional neural network, CNN)在许多图像分类任务中发挥了关键作用^[1]。CNN 通常包括卷积和池化两大操作,卷积操作利用卷积核来捕捉输入数据的特征,而池化操作用于降低数据维度。现有许多研究表明,池化层在提升 CNN 模型性能方面具有不可替代的作用。

为了进一步改善 CNN 模型的性能,研究者们提出了多种方案以改进传统的池化操作。当前池化操作的改进方案大致可以分为两类:基于算子的优化方案和基于权重的优化方案。基于算子的优化方案往往采用新的算子来代替原先的平均池化或最大池化操作,以保留更丰富的信息。例如,文献[2]提出了一种核池化的方法,该方法可在实现降维的同时保留邻域信息,显著提升了 CNN 在场景解析任务中对长距离依赖关系的捕捉能力。文献[3]则基于最优传输理论,设计了一种广义全局池化的方法,其通过从多个维度对输入数据进行降维以保留更精确的信息,进一步提升了模型的分类性能。文献[4]提出了一种随机池化的方法,通过引入随机性来降低模型过拟合的风险,该方法进一步增强了模型的泛化性能。文献[5]基于 Koopman 算子,设计了一种即插即用的参数化池化模块,该方案可保留序列数据中的高阶动态信息。

与上述策略不同,基于权重的优化方案在池化步骤中引入了权重分配机制,通过对池化窗口内部元素实施加权求和操作,实现了更为精细的处理。该方案能够根据不同任务的需求灵活调整权重配置,有效提升了模型的适应性和灵活性。例如,文献[6]提出了一种权重池化的方法,该方法可自适应地为窗口内不同元素分配权重,在多项分类任务上均取得了优异的分类性能。文献[7]在传统池化策略的基础上,引入了动态门控机制来高效整合高价值信息,在多项图分类任务中展现出了卓越的性能。文献[8]构建了一种动态核池化模块,该模块可自适应整合多种参数配置下的关键特征,进一步增强了模型的鲁棒性。

上述改进的池化方法能够进一步提升模型的分类性能,但这些方法因沿袭了滑动窗口机制,依据预设规则选取滑窗内的关键信息而未能充分考虑不同滑窗间的信息对于最终目标的差异化重要性。具体地,在每次池化操作中,非目标区域所含的噪声信息仍然被保留于特征图中,并且随着层级间的不断传播导致噪声累积,致使模型产生不可靠的结论。

针对此问题,将借鉴人类视觉系统处理外界信息过程中关注目标重要特征的行为,设计一种新颖的双重模糊注意力机制(即模糊通道和模糊空间注意力机制)以代替 CNN 中的池化操作,并提出一种基于双重模糊注意力机制的图像分类方法(convolutional neural network with dual fuzzy attention mechanism, DFAM-CNN)。所设计的双重模糊注意力机制将从特征图的通道和空间两个方向上关注目标重要特征信息。为解决从输入层到隐藏层特征映射过程中的不确定性问题,通过引入模糊逻辑技术,模糊通道、

空间注意力机制将分别对各自输入的所有特征图利用模糊隶属度函数进行映射,并确定与原特征图一一对应的重要模糊化特征图;为确定特征图不同通道和不同空间元素的重要性,利用所有重要模糊化特征图,即对应于人类视觉系统处理外界信息时关注目标重要特征的行为,实现所有特征图通道权重和特征图内每个元素权重的计算。具体地,先将所求通道权重对应赋给 CNN 卷积层输出的特征图进行初步强化,随后利用所求元素权重对初步强化输出的特征图进行二次强化,以突出目标在通道和空间方向上有意义的特征。最后通过模糊聚合操作实施特征图的池化,实现提升目标的特征表示能力。本文主要贡献如下:

1)借鉴人类视觉系统在处理外界信息时常会关注目标重要特征和结构,同时弱化非目标区域这一行为,从特征图通道和空间两个方向分配特征图通道权重和特征图内每个元素权重,并形成一种新颖的注意力机制,即双重模糊注意力机制,包括模糊通道注意力机制和模糊空间注意力机制。

2)所提双重模糊注意力机制由于引入了模糊逻辑技术,可有效抑制数据中的噪声在 CNN 模型中的逐层传播。

3)大量实验结果验证了所提双重模糊注意力机制的有效性,且所设计的 DFAM-CNN 与最新方法比较能够取得竞争性的性能。

1 相关工作

1.1 人类视觉系统

人类视觉系统是人类感知世界的主要方式之一,负责捕捉、处理和解释来自外界的视觉信息。人类视觉系统在处理外界信息时表现出高度的选择性和优先级排序能力。当人类观察一个场景或物体时,视觉系统会将注意力集中在目标的重要特征和结构上,这些特征通常包括显著的形状、颜色、纹理或运动等。如图 1 所示,在观察小猫时,视线会集中于小猫所在区域,且距离中心区域越近,越能凸显小猫的详细特征,这些特征在视觉上更为突出和显著。

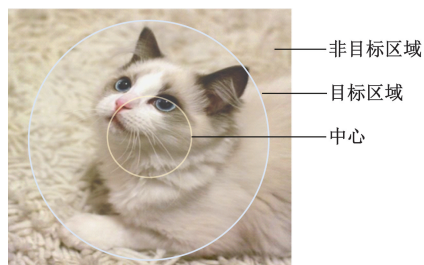


图 1 人类视觉系统示例

Fig. 1 Example of human visual system

同时,人类视觉系统会相对弱化对非目标区域的关注。在处理大量视觉信息时,与目标不相关或不重要的背景信息会被自动过滤掉或降低其重要性。这种选择性关注的能力是人类视觉系统的一个重要特性,能够在复杂的环境中有效地提取和处理信息。因此,本文将模拟人类视觉系统

观察目标的活动过程,从通道和空间两个方向上进行权重计算并赋给原特征图,使原特征图中目标区域信息进一步加强,同时弱化非目标区域。

1.2 注意力机制

注意力机制是一种模仿人类视觉注意力的方法,常与 CNN 结合使用^[9-11]。注意力机制使模型更加关注与目标相关的重要特征和结构,同时弱化非目标区域^[12]。当前,注意力机制主要涵盖三大类别,分别是通道注意力机制、空间注意力机制以及通道与空间的联合注意力机制。

通道注意力机制专注于不同特征通道之间的重要性,通过建模各个特征通道的权重以增强或抑制不同的通道,进而突出对特定任务更为重要的特征。该机制的代表性工作包括 SENet (squeeze-and-excitation networks)^[13] 和 GSoP-Net(global second-order pooling networks)^[14]。空间注意力机制关注于图像中不同空间位置的重要性,通过生成空间特征图来加权特征图中的不同位置,从而突出图像中的关键区域。空间注意力机制的代表模型包括 SA-Net (shuffle attention network)^[15] 和 CCNet (criss-cross network)^[16]。联合注意力机制结合了通道注意力和空间注意力,同时考虑特征通道和空间位置的重要性。该机制能够更全面地处理目标信息,提高模型的性能,代表性工作有双重注意力机制生成对抗网络^[17] 和 CBAM (convolutional block attention module)^[18]。

上述关于注意力机制的方法已被证明能够提高基于 CNN 模型的分类精度,在处理视觉任务时更加高效和准确。特征图之间的相关性对于提高 CNN 模型在图像分类任务中的性能至关重要。因此,本文同时考虑通道注意力机制和空间注意力机制以解决特征图之间的通道相关性和空间相关性问题。

1.3 模糊逻辑技术

模糊逻辑技术是一种处理模糊和不确定性信息的数学方法。将模糊逻辑技术引入到 CNN 池化操作中可以有效降低噪声在 CNN 模型中的逐层传播,进而能够准确提取特征。文献[19]提出了一种基于卷积神经网络的多焦点图像融合方法,该方法利用模糊集理论来描述多焦点图像融合过程中的模糊性特征信息,同时采用多尺度分析技术提升了图像融合的效果。文献[20]提出一种卷积模糊神经网络,采用自适应模糊逻辑对高频子带进行融合,有效提升了多焦点图像融合的训练速度和效率。文献[21]提出一种基于模糊集的模糊池化操作,通过模糊化、特征图邻域的聚合和去模糊化以构建模糊池化层,从而能够较好地处理特征图中的局部不确定性。

将模糊逻辑技术引入到 CNN 模型的方法能够增强 CNN 的鲁棒性和泛化能力。从输入层到隐藏层的特征映射过程中的不确定性问题可能导致图像中的噪声在 CNN 模型中逐层传播,进而影响特征表示能力。因此,本文考虑通过引入模糊逻辑技术并与通道、空间注意力机制结合,分别提出模糊通道注意力机制和模糊空间注意力机制以改善 CNN 模型性能。

2 双重模糊注意力机制模型

本节将深入探讨基于双重模糊注意力机制的图像分类方法 DFAM-CNN,包括其构成要素和计算方法。首先概述了整体框架,随后详细阐述了模糊映射的计算步骤,并进一步介绍了基于模糊逻辑技术所构建的模糊通道注意力机制和模糊空间注意力机制的计算步骤。

2.1 总体框架

本文提出了一种基于双重模糊注意力机制的图像分类方法 DFAM-CNN,其总体架构如图 2 所示。

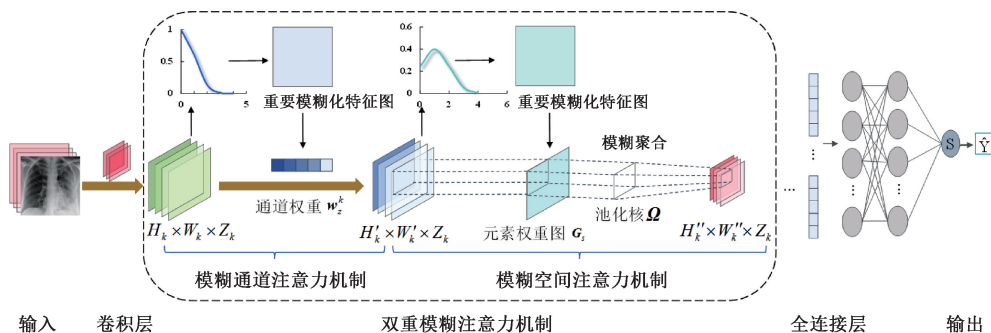


图 2 DFAM-CNN 的总体架构

Fig. 2 The overall architecture of DFAM-CNN

在基于 CNN 的图像分类算法框架内,池化操作作为一种高效的降维技术,得到了广泛的应用。该操作通过预设的滑动模式,选择性地保留所有窗口内的最大值/均值,以此实现特征降维。然而,这一做法隐含地假设了图像各区域内所有信息的重要性是一致的。受人类视觉系统在处理外界信息时优先关注目标重要特征和结构的启发,并

借鉴这一原理,本文利用双重模糊注意力机制代替 CNN 模型的池化层以增强图像重要区域的关键特征表征,从而提高 CNN 模型的分类性能,其余架构均与 CNN 模型保持一致。

当输入给定图像数据时,经过卷积操作并将 CNN 卷积层输出的特征图记为 $H_k \times W_k \times Z_k$,其中 k 代表 CNN

第 k 个卷积层, H_k 代表第 k 个卷积层中特征图的高度, W_k 代表第 k 个卷积层中特征图的宽度, Z_k 代表第 k 个卷积层中特征图的通道数, 模糊通道、空间注意力机制将分别对各自输入的所有特征图利用模糊隶属度函数进行映射, 并确定与原特征图一一对应的重要模糊化特征图, 利用所有重要模糊化特征图分别计算通道权重 w_z^k 和元素权重图 G_s 。

在模糊通道注意力机制中, 通道权重 w_z^k 对应赋给 CNN 卷积层输出的特征图 $H_k \times W_k \times Z_k$ 以得到初步强化输出的特征图 $H'_k \times W'_k \times Z_k$ 。在模糊空间注意力机制中, 根据所求元素权重图 G_s 和池化核 Ω 进行模糊聚合操作以实施特征图的池化, 得到二次强化输出的特征图 $H''_k \times W''_k \times Z_k$, 并将此结果作为池化层的输出, 其大小由特征图的窗口大小和池化窗口大小共同决定。

2.2 模糊映射

模糊逻辑技术通过模糊映射将输入映射到新的空间, 从而实现对不确定性问题的建模和处理。模糊映射过程首先要确定模糊划分数 P 和模糊隶属度函数, 实验中选取的模糊划分数 P 将在较小的奇数 $\{3, 5, 7, 9\}$ 中进行搜索, 同时选择高斯函数作为模糊隶属度函数^[22]。在确定最优模糊划分数时, 实验结果表明模糊划分数设置为 3 时 DFAM-CNN 在多数情况下能够取得最好的性能。此外, 模糊划分数 P 与高斯模糊隶属度函数个数一致, 即每个模糊集合对应一个高斯模糊隶属度函数。

基于模糊划分数 P 生成高斯模糊隶属度函数的两个参数, 即中心 $c_{p,z}^k$ 和核宽度 $\sigma_{p,z}^k$ 。将高斯模糊隶属度函数的核宽度 $\sigma_{p,z}^k$ 统一设置为 1 时可在多数情况下获得最优性能的同时简化双重模糊注意力机制的操作流程。此外, 高斯模糊隶属度函数的中心 $c_{p,z}^k$ 根据特征图内元素的最大值

和最小值进行确定, 计算公式为:

$$c_{p,z}^k = \frac{x_{\max,z}^k - x_{\min,z}^k}{P-1}(p-1), \quad (1)$$

$$1 \leq p \leq P, 1 \leq z \leq Z_k$$

其中, $x_{\max,z}^k$ 和 $x_{\min,z}^k$ 分别代表第 k 个卷积层中第 z 个特征图中所有元素的最大值与最小值, P 代表模糊划分数, $1 \leq p \leq P$, z 代表第 k 个卷积层中的第 z 个特征图, $1 \leq z \leq Z_k$ 。利用高斯模糊隶属度函数对第 k 个卷积层输出的所有特征图中所有元素进行模糊化, 即:

$$\mu^p(x_{ij,z}^k) = \exp\left(-\frac{(x_{ij,z}^k - c_{p,z}^k)^2}{2(\sigma_{p,z}^k)^2}\right) \quad (2)$$

其中, $x_{ij,z}^k$ 代表所述第 k 个卷积层中的第 z 个特征图中的第 i 行、第 j 列元素, $\mu^p(x_{ij,z}^k)$ 代表与元素 $x_{ij,z}^k$ 对应的第 p 个高斯函数的模糊隶属度值, 根据所求模糊隶属度值生成模糊化特征图。每个高斯模糊隶属度函数对应一个模糊化特征图, 即 CNN 卷积层输出的第 z 个特征图采用 P 个高斯模糊隶属度函数进行映射, 将生成 P 个模糊化特征图。

2.3 模糊通道注意力机制

在模糊通道注意力机制中, 针对 CNN 卷积层输出的特征图 $H_k \times W_k \times Z_k$, 通过引入模糊逻辑技术对 CNN 卷积层输出的特征图利用高斯模糊隶属度函数进行模糊映射, 并确定与特征图一一对应的重要模糊化特征图, 利用所有重要模糊化特征图计算与 CNN 卷积层输出的特征图对应的通道权重 w_z^k , 其中 k 代表卷积神经网络中的第 k 个卷积层, z 代表所有特征图中第 z 个特征图, 将所求通道权重 w_z^k 对应赋给 CNN 卷积层输出的特征图以得到初步强化输出的特征图 $H'_k \times W'_k \times Z_k$, 通道权重的计算步骤如图 3 所示。

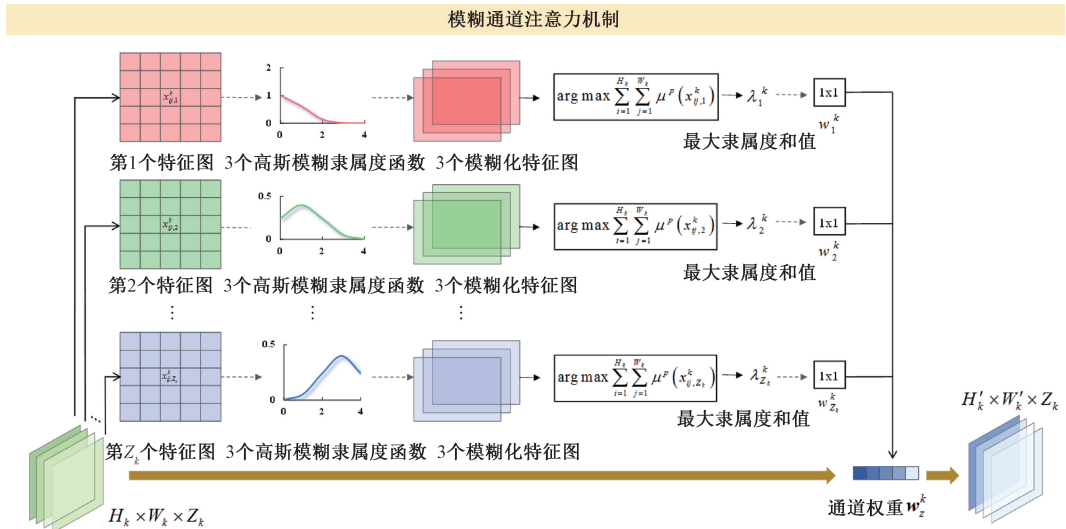


图 3 模糊通道注意力机制

Fig. 3 Fuzzy channel attention mechanism

如图 3 所示,最左侧为 CNN 卷积层输出的特征图 $H_k \times W_k \times Z_k$,利用 3 个高斯模糊隶属度函数对每张特征图进行模糊映射并生成 3 个模糊化特征图,将这些模糊化特征图中所有元素各自求和并比较大小,将比较后的最大值记为最大隶属度和值。特别地,根据最大隶属度和值确定与 CNN 卷积层输出的特征图一一对应的重要模糊化特征图。第 k 个卷积层中第 z 个特征图对应的最大隶属度和值 λ_z^k 计算公式为:

$$\lambda_z^k = \arg \max \sum_{i=1}^{H_k} \sum_{j=1}^{W_k} \mu^p(x_{ij,z}^k), \quad (3)$$

$$1 \leq p \leq P, 1 \leq z \leq Z_k$$

其中, $x_{ij,z}^k$ 代表所述第 k 个卷积层中第 z 个特征图中的第 i 行、第 j 列元素, $\mu^p(x_{ij,z}^k)$ 代表与元素 $x_{ij,z}^k$ 对应的第 p 个高斯函数的模糊隶属度值。每个特征图对应一个最大隶属度和值以及重要模糊化特征图,根据所求最大隶属度和值计算与 CNN 卷积层输出的特征图对应的通道权重

重 w_z^k ,第 k 个卷积层中第 z 个特征图对应的通道权重计算公式为:

$$w_z^k = \frac{\lambda_z^k}{\lambda_1^k + \lambda_2^k + \dots + \lambda_{Z_k}^k}, 1 \leq z \leq Z_k \quad (4)$$

其中, w_z^k 代表第 k 个卷积层中第 z 个特征图对应的通道权重, λ_z^k 代表第 k 个卷积层中第 z 个特征图对应的最大隶属度和值。将所求通道权重 w_z^k 对应赋给 CNN 卷积层输出的特征图以得到初步强化输出的特征图 $H'_k \times W'_k \times Z_k$ 。

2.4 模糊空间注意力机制

在模糊空间注意力机制中,针对初步强化输出的特征图 $H'_k \times W'_k \times Z_k$,与模糊通道注意力机制中模糊映射过程一致,利用高斯模糊隶属度函数进行模糊映射并确定最大隶属度和值以及重要模糊化特征图,利用所有重要模糊化特征图计算特征图内每个元素权重,并形成一张元素权重图 G_s ,元素权重的计算步骤如图 4 所示。

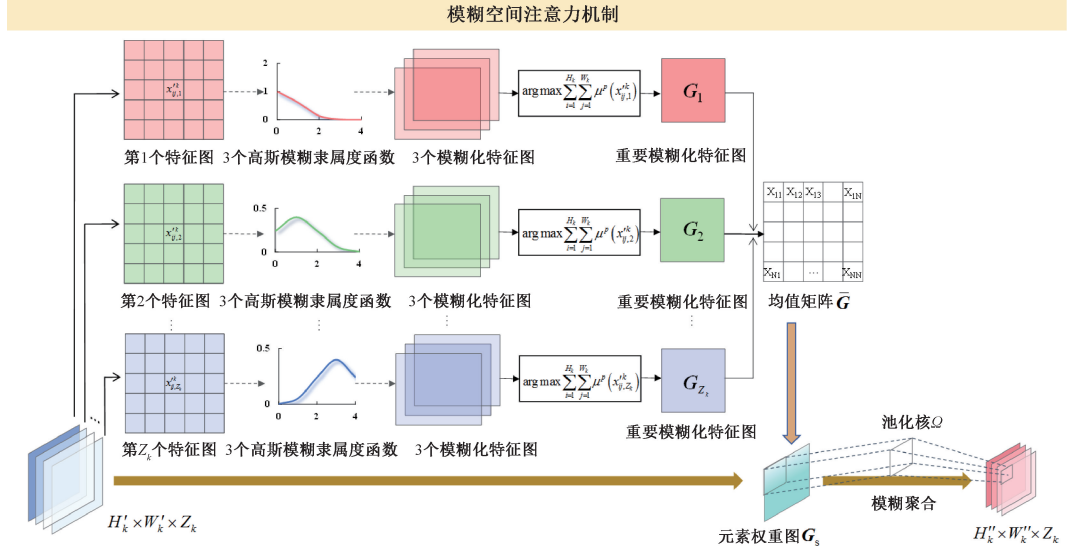


图 4 模糊空间注意力机制

Fig. 4 Fuzzy spatial attention mechanism

如图 4 所示,最左边为初步强化输出的特征图 $H'_k \times W'_k \times Z_k$,即模糊通道注意力机制的输出。利用 3 个高斯模糊隶属度函数进行模糊映射,并确定与初步强化输出的特征图一一对应的重要模糊化特征图,将所有重要模糊化特征图在通道方向上求平均,记为均值矩阵 \bar{G} ,其大小为 $H_k \times W_k$,计算如下:

$$\bar{G} = \frac{G_1 + G_2 + \dots + G_{Z_k}}{Z_k}, 1 \leq z \leq Z_k \quad (5)$$

其中, G_z 代表第 z 个特征图对应的重要模糊化特征图, z 代表第 k 个卷积层中第 z 个特征图, $1 \leq z \leq Z_k$ 。

利用所求 Z_k 个重要模糊化特征图在通道方向上求平均得到均值矩阵 \bar{G} 。将均值矩阵中每个位置元素除以均

值矩阵内所有元素加和,以求得特征图内每个元素权重,并形成一张元素权重图 G_s ,其大小为 $H_k \times W_k$,计算如下:

$$G_s = \frac{x_{ij,\bar{G}}}{\sum_{i=1}^{H_k} \sum_{j=1}^{W_k} x_{ij,\bar{G}}} \quad (6)$$

其中, $x_{ij,\bar{G}}$ 代表均值矩阵中第 i 行和第 j 列元素的值, $\sum_{i=1}^{H_k} \sum_{j=1}^{W_k} x_{ij,\bar{G}}$ 代表均值矩阵内所有元素加和。

根据所求元素权重图 G_s 和池化核 Ω 进行模糊聚合操作以实施特征图的池化。针对初步强化输出的特征图 $H'_k \times W'_k \times Z_k$ 和元素权重图 G_s ,滑动池化窗口以实现下采样,池化窗口同时在初步强化输出的每个特征图和元素权重图上移动,并将池化窗口范围内的每个特征图中元素与

元素权重图中元素进行点积并求和,将求和结果作为移动一次池化窗口后的池化结果。初步强化输出的第 z 个特征图的池化结果计算公式为:

$$\mathbf{Y}_{i,j} = \sum_{u=1}^n \sum_{v=1}^n (\boldsymbol{\Omega}_{u,v} \cdot \mathbf{X}_{(i+u-1,j+v-1)}^z \cdot \mathbf{G}_{s(i+u-1,j+v-1)}) \quad (7)$$

其中, $\mathbf{Y}_{i,j}$ 代表第 z 个特征图池化结果中第 i 行和第 j 列元素的值,池化核记为 $\boldsymbol{\Omega}$, 大小为 $n \times n$, $\mathbf{X}_{(i+u-1,j+v-1)}^z$ 代表第 z 个特征图在池化窗口范围内的元素, 大小为 $H_k \times W_k$, $\mathbf{G}_{s(i+u-1,j+v-1)}$ 代表元素权重图在池化窗口范围内的元素, 大小为 $H_k \times W_k$ 。

根据此方式依次同时在每个初步强化输出特征图和元素权重图上移动池化窗口,最终得到第 k 个卷积层输出的二次强化的所有特征图 $H_k'' \times W_k'' \times Z_k$, 并作为第 k 个卷积层之后的池化层输出,其大小由特征图的窗口大小和池化窗口大小共同决定。

3 实验及结果分析

CNN 的性能在很大程度上取决于其网络结构的层数

和卷积核的大小。因此,本文选择以 LeNet-5 模型为基本框架进行实验,以便有效地验证所提双重模糊注意力机制对 DFAM-CNN 模型性能的影响。具体地,将利用双重模糊注意力机制代替 LeNet-5 的池化层,设计一种基于双重模糊注意力机制的图像分类方法 DFAM-CNN,深入探究模糊通道和空间注意力机制在降低噪声方面的效果,并分析其对分类性能的影响。首先展示了 DFAM-CNN 在 MedMNIST 数据上的分类性能,随后通过对比实验和消融实验验证了双重模糊注意力机制在提升分类性能方面的有效性,最后进行可视化分析展示了 DFAM-CNN 在降低噪声方面的效果。

3.1 数据集

本文采用上海交通大学公开提供的医学图像数据集 MedMNIST^[23]进行实验。该数据集由 10 个预处理数据集组成,涵盖不同数据模式(如 X 射线、OCT、超声、CT),不同分类任务(二进制/多类)和数据大小(从 100~100 000)。MedMNIST 数据集详细信息如表 1 所示。

表 1 MedMNIST 数据集信息
Table 1 Information of the MedMNIST dataset

名称	大小	类别	训练集	测试集
BreastMNIST(BRE)	28×28×1	2	546	156
ChestMNIST(CHE)	28×28×1	2	78 468	22 433
DermaMNIST(DER)	28×28×3	7	7 007	2 005
OCTMNIST(OCT)	28×28×1	4	97 477	1 000
OrganMNIST_Axial(AXI)	28×28×1	11	34 581	17 778
OrganMNIST_Coronal(COR)	28×28×1	11	13 000	8 268
OrganMNIST_Sagittal(SAG)	28×28×1	11	13 940	8 829
PathMNIST(PAT)	28×28×3	9	89 996	7 180
PneumoniaMNIST(PNE)	28×28×1	2	4 708	624
RetinaMNIST(RET)	28×28×3	5	1 080	400

3.2 对比算法和参数设置

在本文实验中,采用了与 LeNet-5 相同的基本架构来设计对比算法,以确保评估过程的客观性和公平性。所选对比算法包括最大池化(Max pooling),平均池化(Average pooling)^[24],参数随机分配的模糊卷积神经网络(ConvFNN)^[25],模糊池化(Fuzzy pooling),基于模糊池化的卷积神经网络图像分类(FP-CNN)^[26],多类模糊加权 AdaBoost(FWAdaBoost)^[27]和多层进化模糊神经网络(MEFNN)^[28]。

在本文中,为确保不同模型之间性能比较的公正性,所有模型均采用统一的超参数配置。具体而言,最大迭代次数被设置为 200,学习率被设置为 0.001,池化核大小则被设置为 2×2。对于本文所提出的方法,模糊划分分数 P 被设定为 3。所有对比算法均使用大小为 28×28 的原始图像进行训练,没有进行数据预处理或增强。所有实验均在

MATLAB 2022b 和配备 Intel(R) Xeon(R) E5-2680v4 2.40 GHz CPU 和 80 G 内存的计算机上实现。

3.3 评价指标

为了评估 DFAM-CNN 的分类性能,采用了 5 个常用分类指标,分别为准确率(ACC)、曲线下面积(AUC)、召回率(Recall)、特异性(SPE)和 F1 评分(F1score),以确保评估的全面性。5 种评价指标的值均在 0~1 之间,值越大表明分类效果越好。计算公式分别如下:

$$ACC = \frac{TP + TN}{TP + TN + FP + FN} \quad (8)$$

$$AUC = P(P_{\text{Positive}} > P_{\text{Negative}}) \quad (9)$$

$$Recall = \frac{TP}{TP + FN} \quad (10)$$

$$SPE = \frac{TN}{TN + FP} \quad (11)$$

$$Precision = \frac{TP}{TP + FP}$$

(12)

$$F1score = 2 \times \frac{Precision \times Recall}{Precision + Recall}$$

(13)

3.4 分类性能

本节实验旨在探究 DFAM-CNN 的分类性能。利用

所设计的双重模糊注意力机制代替 LeNet-5 架构中的池化层,其余架构均与 LeNet-5 保持一致。在 MedMNIST 数据集上采用五个常用的评价指标来衡量所提模型的性能。实验结果均在相同参数设置下运行算法 10 次后取最大值。表 2 展示了 DFAM-CNN 在 MedMNIST 数据集上的分类性能。

表 2 DFAM-CNN 在 MedMNIST 数据集上分类性能
Table 2 Classification performance of DFAM-CNN on the MedMNIST dataset

数据	ACC	AUC	Recall	SPE	F1
BreastMNIST(BRE)	83.08	86.84	75.34	70.00	85.22
ChestMNIST(CHE)	89.31	61.35	76.43	70.00	67.15
DermaMNIST(DER)	76.88	70.17	60.24	85.71	41.45
OCTMNIST(OCT)	65.00	83.66	81.43	79.00	60.84
OrganMNIST_Axial(AXI)	60.24	91.14	55.76	95.95	55.65
OrganMNIST_Coronal(COR)	89.94	86.03	84.13	94.77	85.63
OrganMNIST_Sagittal(SAG)	80.08	78.10	76.34	93.49	73.59
PathMNIST(PAT)	77.17	90.23	85.54	88.89	73.26
PneumoniaMNIST(PNE)	92.50	91.09	81.34	78.36	78.46
RetinaMNIST(RET)	83.50	79.17	54.88	87.98	62.13

注:表 2 中展示结果均为 10 次实验中最大结果。

3.5 对比实验

本节实验旨在深入研究双重模糊注意力机制在处理通道相关性和空间相关性问题上的效能,并探究其是否能够显著提升模型分类性能。对比了所提模型 DFAM-CNN 与 7 种改进池化操作方

法。为了确保结果的可靠性和可重复性,所有算法均运行了 10 次,最终展示的 ACC 是基于这 10 次运行结果计算得出的平均值。具体实验结果如表 3 所示,其中,最佳结果以黑色加粗字体标出。

表 3 所有对比算法在 MedMNIST 数据集上分类性能(ACC/%)

Table 3 Classification performance of all comparison algorithms on the MedMNIST dataset (ACC/%)

方法	BRE	CHE	DER	OCT	AXI	COR	SAG	PAT	PNE	RET
Max pooling	73.08	89.11	66.88	45.21	62.83	59.21	42.88	57.17	79.07	43.19
Average pooling	73.08	88.21	66.88	45.00	60.78	54.81	41.31	59.25	62.50	43.50
ConvFNN	76.28	89.16	71.57	59.00	80.62	79.29	63.65	74.79	82.37	48.50
Fuzzy pooling	73.08	89.71	66.88	35.56	37.08	37.77	29.45	75.39	62.64	42.38
FP-CNN	80.13	89.17	71.37	64.60	84.80	84.85	69.53	79.83	83.55	52.33
FWAdaBoost	80.77	87.49	70.82	49.70	30.35	30.62	30.15	43.38	83.65	50.25
MEFNN	75.00	87.53	67.88	47.40	52.67	45.63	34.53	67.65	60.12	44.25
DFAM-CNN	81.75	89.21	74.28	65.00	65.24	85.72	77.47	77.06	90.73	63.09

注:表 3 中展示结果均为 10 次实验后的平均结果。

1)池化层的主要作用是减少参数量和计算复杂度。通过下采样操作,池化层可以降低特征图的空间维度,减少后续层的计算负担,同时保留重要的特征信息。与最大池化 Max pooling 和平均池化 Average pooling 两种传统的池化方法相比,DFAM-CNN 虽然增加了通道权重和空间位置权重的参数量,但通过两次权重调整机制,能够更有效地提取目标特征信息。这种改进使得 DFAM-CNN 在 OCT、COR、SAG 和 PNR 数据集上的 ACC 得到了显著

提升,验证了所提双重模糊注意力机制的有效性。此外,从表 3 中可以看出,DFAM-CNN 的 ACC 在大多数情况下优于最大池化 Max pooling 和平均池化 Average pooling,这可能是因为 DFAM-CNN 充分考虑了特征的通道相关性和空间相关性,从而展现出较强的特征表示能力。

2)BRE 数据集和 RET 数据训练量分别为 546 和 1 080,属于小规模数据集,DFAM-CNN 在这两个数据上取得最优 ACC,分别为 81.75%和 63.09%。OCT 和 PAT

数据训练量分别为 97 477 和 89 996,属于较大规模数据集,DFAM-CNN 在 OCT 数据集上取得最优 ACC 为 65.00%,在 PAT 数据集上,DFAM-CNN 的 ACC 低于 FP-CNN 方法 2.77%,但超出 Max pooling 和 Average pooling 近 20 个百分点,在 ACC 上仍然取得了较大提升。由此表明,双重模糊注意力机制能够准确提取特征,DFAM-CNN 在中小规模数据集上具有较强的适应性。

3)与同样采用模糊逻辑技术和模糊聚合操作的 Fuzzy pooling 相比,DFAM-CNN 不仅解决了特征图中的局部不确定性,而且考虑了通道之间的相关性。表 3 中的结果显示,在 MedMNIST 数据集上,DFAM-CNN 的 ACC 在多数情况下均优于 Fuzzy pooling,这表明双重模糊注意力机制能够有效增强模型对目标特征的表征能力。

3.6 消融实验

为系统评估关键超参数对模型分类性能的影响,本文

详细分析了学习率和模糊划分数在取不同值时,DFAM-CNN 在 OCT、AXI、COR 和 PNE 4 种数据集上的分类性能。其中,学习率将在 $\{10^{-3}, 10^{-2}, 10^{-1}\}$ 范围内搜索,模糊划分数在 $\{3, 5, 7, 9\}$ 中搜索,具体结果如图 5 所示。其中,图 5(a)~(d)分别对应 OCT、AXI、COR 和 PNE 四种数据集在不同超参数设置下的 ACC 变化情况。由图 5 可观察到,当学习率恒定时,模型在四种数据集上的分类性能均随着模糊划分数值的增大而减小或不变。这可能是因为,模糊划分数值越大,搜索范围越大,引入的不确定性因素也会增多。这些不确定性因素的增多,会在一定程度上干扰模型的决策过程,导致模型的分类性能变得不够稳定,进而影响了其在各种数据集上的表现。此外,当模糊划分数值一定时,学习率取 0.001 可在多数情况下可以取得最优。因此,本文中学习率和模糊划分数被设置为 0.001 和 3。

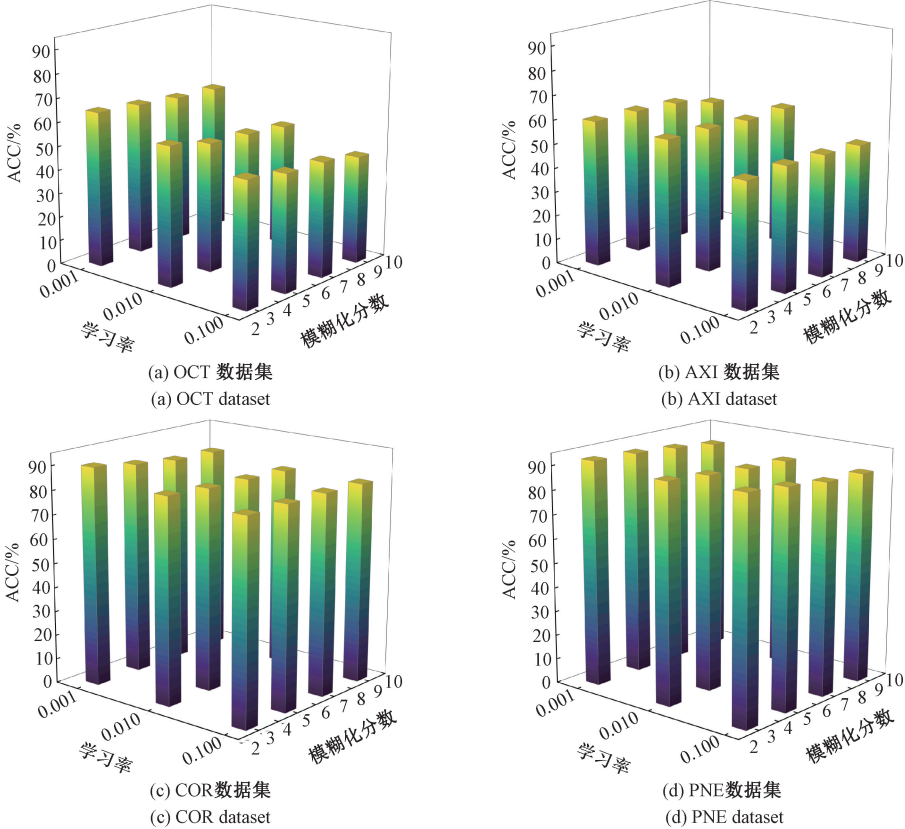


图 5 超参数学习率和模糊划分数对分类性能的影响

Fig. 5 The influence of hyperparameter learning rate and fuzzy partition on classification performance

除了对超参数学习率和模糊划分数进行深入的讨论分析,为进一步验证所提出的两种特征融合机制——模糊通道注意力机制与模糊空间注意力机制,是否有助于提升模型的性能,本节设计了两变体网络。具体而言,以 LeNet-5 作为基准模型,分别将模糊通道注意力机制和模糊空间注意力机制与 LeNet-5 结合,以此单独评估每种模糊注意力机制对分类性能的影响。图 6 展示了消融实

验中结合模糊通道注意力机制和模糊空间注意力机制的 ACC 变化情况。具体而言,图 6(a)~(d)分别展示了不同网络模型在 OCT、AXI、COR 和 PNE 数据集上的 ACC 变化情况。

通过消融实验探究在 LeNet-5 基础上引入模糊注意力机制对 CNN 模型性能的影响。图 6(a)~(d)中消融实验结果表明,基于 4 种不同规模数据,DFAM-CNN 在 OCT、

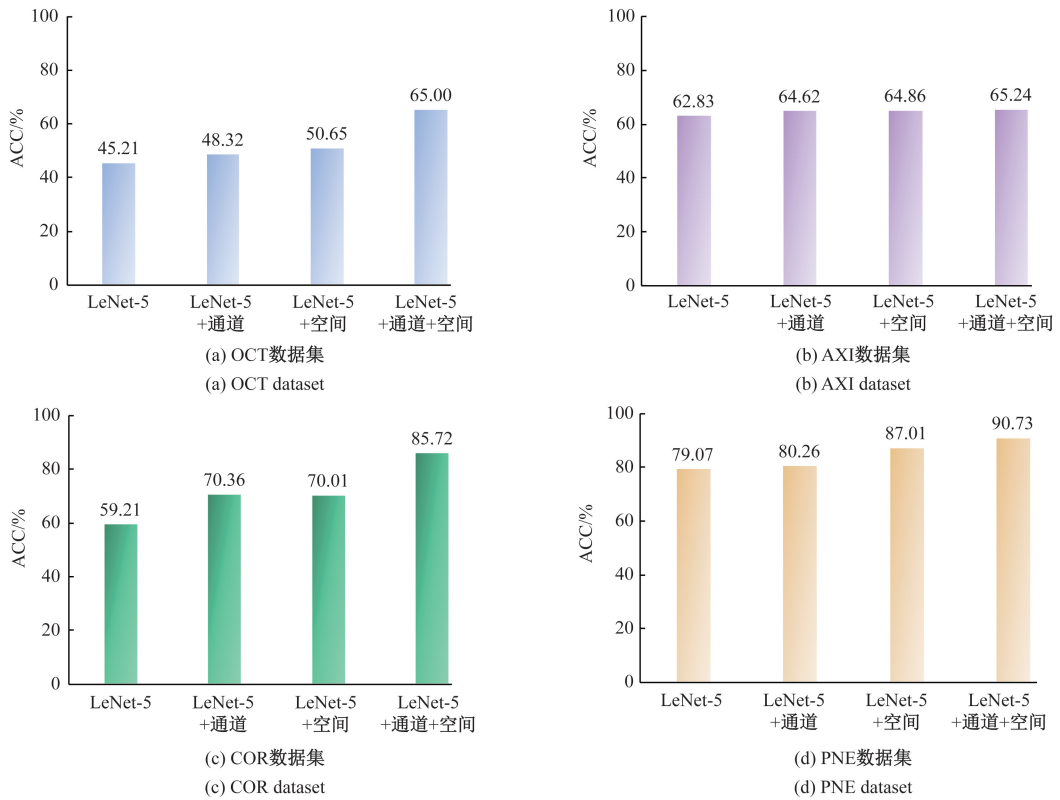


图6 在OCT、AXI、COR和PNE数据集上的消融实验结果

Fig. 6 Results of ablation experiments on OCT, AXI, COR, and PNE datasets

AXI、COR和PNE数据上的ACC均有提升,分别提升了19.79%、2.41%、26.51%和11.66%。

特别地,在OCT数据集上,首先将基准模型LeNet-5与模糊通道注意力机制结合,ACC从45.21%提升至48.32%,表明模糊通道注意力机制能够用于解决特征图之间的相关性,有利于提升分类精度。其次,将基准模型LeNet-5与模糊空间注意力机制结合,ACC提升了5.44%,由此表明模糊空间注意力机制能够用于准确提取特征图的内部特征与其相邻特征。最后以LeNet-5为基准模型,同时结合模糊通道注意力机制和模糊空间注意力机制,ACC提升了19.79%,由此表明DFAM-CNN有助于解决特征图之间的相关性包括通道相关性和空间相关性,进而提升分类性能。

3.7 降噪效果可视化

为了进一步探究DFAM-CNN在降低噪声方面的效果,在测试集中添加不同强度的高斯噪声进行验证。具体如下,在测试集中添加不同强度的高斯噪声,以模拟图像在采集和传输过程中可能遇到的噪声干扰。高斯噪声是通过在图像各像素点上叠加符合正态分布的随机扰动实现的,噪声强度通常由标准差这一关键参数表征,标准差量化了噪声的离散程度,其数值与噪声的波动幅度呈正相关,即标准差越大,表明噪声分布范围越广,对图像造成的干扰程度也越显著。

本节实验中引入了不同强度的高斯噪声,以比较和分析噪声强度对分类性能的具体影响,噪声强度分别为0、0.05、0.15和0.25。具体而言,如图7所示,第1行为无噪声,第2行噪声强度为0.05,第3行噪声强度为0.15,第4行噪声强度为0.25。图7详细展示了在图像中添加不同强度的高斯噪声后,LeNet-5与DFAM-CNN两种模型的热力图可视化结果。从图7中可以看出,随着噪声强度的逐渐增强,原始图像愈发模糊,同时两个模型的热力图均展现出一定程度的变动。值得注意的是,DFAM-CNN模型的热力图在图像轮廓方面始终能够保持相对清晰的识别能力,而LeNet-5模型则随噪声强度的增大表现出愈发混乱的趋势,最终甚至出现紊乱现象。这表明,本文所提出的DFAM-CNN模型能够有效抑制噪声的干扰,从而验证了模糊注意力策略的有效性。

同时,本节还深入探究了在随机添加不同强度高斯噪声的条件下,模型在PNE及OCT数据集上的ACC性能变化,具体结果如图8(a)、(b)所示。实验结果表明,DFAM-CNN在添加了不同强度高斯噪声后的测试集上仍保持了较高的分类ACC,显示出了良好的降低噪声能力。具体地,高斯噪声强度在0.05和0.15时,LeNet-5与DFAM-CNN的ACC均略有下降,但DFAM-CNN的ACC仍然保持在较高水平,而在添加强度为0.25的高斯噪声时,DFAM-CNN的ACC下降更为明显,但仍优于传统的



图 7 两种模型在不同强度高斯噪声条件下的热力图可视化效果

Fig. 7 Visualization effects of the heat map of two models under varying intensities of Gaussian noise

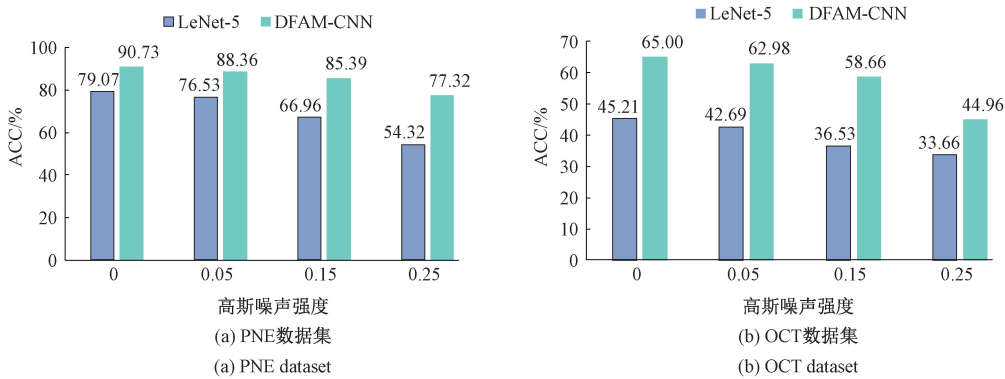


图 8 在 PNE 和 OCT 数据集上添加不同强度高斯噪声的性能变化

Fig. 8 Performance changes of adding varying intensities of Gaussian noise on PNE and OCT datasets

LeNet-5 模型。这表明,所提出的双重模糊注意力机制能够有效提取图像中的关键特征信息,同时显著抑制噪声干扰,从而提升了模型的鲁棒性。

3.8 应用实例

为了验证 DFAM-CNN 的广泛适用性,选用了两个经典的基准数据集 MNIST^[29] 和 Fashion-MNIST^[30] 作为应用实例进行实验。图 9(a)、(b)展示了来自 MNIST 和 Fashion-MNIST 数据集的部分图像示例。MNIST 数据集包含 70 000 张 28×28 像素的灰度图像,涵盖了 0~9 的手写数字,其中 60 000 张用于训练,10 000 张用于测试。Fashion-MNIST 数据集在图像大小、数据格式以及训练和

测试集划分结构上与 MNIST 保持一致,但内容为服装和配饰,而非手写数字。

表 4 展示了 MNIST 和 Fashion-MNIST 数据集详细信息。本节实验中对比较算法、参数设置以及评价指标均与上述 MedMNIST 数据实验保持一致。

表 4 MNIST 和 Fashion-MNIST 数据集信息

Table 4 Information of MNIST and Fashion-MNIST datasets				
名称	大小	类别	训练集	测试集
MNIST	28×28×1	10	60 000	10 000
Fashion-MNIST	28×28×1	10	60 000	10 000

表 5 展示了所有对比算法在 MNIST 和 Fashion-MNIST 数据集上的分类性能。由表 5 可以看出,所提模型 DFAM-CNN 在两个应用实例数据集上均取得了最优性能。具体地,在 MNIST 手写体数据集上,DFAM-CNN 取得了 98.82% 的 ACC,超越了传统池化方法(Max pooling 为 98.20%,Average pooling 为 96.52%),这充分验证了双重模糊注意力机制在特征权重方面的有效性,能够在通道和空间方向上突出与目标相关的特征,进而准确提取特征。在更具挑战性的 Fashion-MNIST 数据集上,面对复杂的服装纹理和结构特征,DFAM-CNN 以 88.60% 的准确率继续保持领先优势,较性能次优的 Max pooling 方

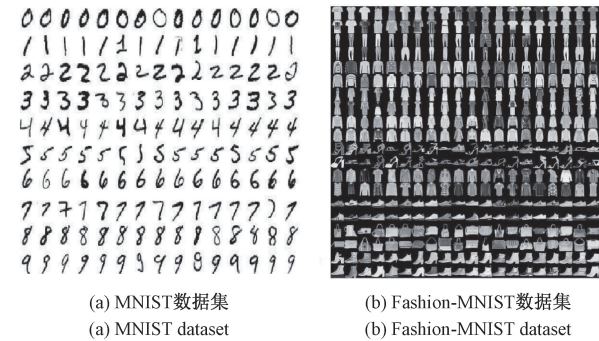


图 9 MNIST 和 Fashion-MNIST 数据集图像示例

Fig. 9 Image examples of MNIST and Fashion-MNIST datasets

法(88.27%)提升了 0.33%。这一结果也有力证明了所提双重模糊注意力机制能够聚焦关键特征区域,同时弱化非目标区域,在复杂图像分类任务中具有一定的竞争力。

表 5 所有对比算法在 MNIST 和 Fashion-MNIST 数据集上分类性能(ACC/%)

Table 5 Classification performance of all comparison algorithms on MNIST and Fashion-MNIST datasets(ACC/%)

方法	MNIST	Fashion-MNIST
Max pooling	98.20	88.27
Average pooling	96.52	87.45
ConvFNN	94.68	80.09
Fuzzy pooling	98.07	87.33
FP-CNN	98.74	87.77
FWAdaBoost	97.89	87.86
MEFNN	97.12	84.68
DFAM-CNN	98.82	88.60

注:表 5 中展示结果均为 10 次实验后的平均结果。

4 结 论

本文提出了一种基于双重模糊注意力机制的图像分类方法。通过引入模糊映射技术以增强目标关键信息的表征,同时弱化非目标区域。利用模糊通道注意力机制和模糊空间注意力机制分别计算通道权重和空间元素权重,以突出图片中感兴趣区域的特征,从而提高模型的泛化能力。将双重模糊注意力机制与 CNN 相结合,所提模型 DFAM-CNN 充分考虑并解决了特征图的通道相关性和空间相关性问题,同时在降低图像噪声方面也具有一定效果。

虽然本文所述方法在许多任务上都取得了一些更优异的性能,但都是在中小规模数据集上进行了测试。未来,拟对自适应的模糊逻辑技术进行研究,实现一种由数据自身分布特点决定的模糊划分确定方法,以便与 GPU 的并行处理能力相结合,使所提双重模糊注意力机制具有处理大规模图像数据集的能力,如 ImageNet 数据集。此外,将进一步研究所提双重模糊注意力机制的性能和特点,使其能够与更复杂的模型(如 AlexNet、GoogLeNet 和残差网络等)相结合,以适应更深层次的图像特征学习。

参考文献

[1] LI Z, LIU F, YANG W, et al. A survey of convolutional neural networks: analysis, applications, and prospects [J]. IEEE Transactions on Neural Networks and Learning Systems, 2021, 33 (12): 6999-7019.

[2] HOU Q, ZHANG L, CHENG M M, et al. Strip pooling: Rethinking spatial pooling for scene parsing[C].

IEEE Conference on Computer Vision and Pattern Recognition, 2020: 4003-4012.

[3] XU H, CHENG M J. Regularized optimal transport layers for generalized global pooling operations [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2023, 45(12): 15426-15444.

[4] WANG SH H, LYU Y D, SUI Y X, et al. Alcoholism detection by data augmentation and convolutional neural network with stochastic pooling[J]. Journal of Medical Systems, 2018, 42(2): 1-11.

[5] WANG X H, XU X, MU Y D. Neural Koopman pooling: Control-inspired temporal dynamics encoding for skeleton-based action recognition [C]. IEEE Conference on Computer Vision and Pattern Recognition, 2023: 10597-10607.

[6] EGGEN K, FAN H C. WAPooling: An adaptive plug-and-play module for feature aggregation in point cloud classification networks[J]. International Journal of Applied Earth Observation and Geoinformation, 2025, 138: 104439.

[7] WANG Z D, YANG L, CHEN T X, et al. Clustering matrix regularization guided hierarchical graph pooling[J]. Knowledge-Based Systems, 2025, 311: 113108.

[8] BASHIVAN P, BAYAT R, IBRAHIM A, et al. Learning adversarially robust kernel ensembles with kernel average pooling [J]. Expert Systems with Applications, 2025, 266: 126017.

[9] 朱江, 杜瑞, 李建奇, 等. 基于注意力机制的曲轴瓦盖上料机器人视觉定位和检测方法[J]. 仪器仪表学报, 2021, 42(5): 140-150.

ZHU J, DU R, LI J Q, et al. Attention mechanism-based vision localization and detection method for crankshaft shingle cover loading robot [J]. Chinese Journal of Scientific Instrument, 2021, 42 (5): 140-150.

[10] 刘秋月, 刘雪峰, 孙绍华. 基于阴影增强和注意力机制的高光谱图像分类[J]. 电子测量技术, 2024, 46(8): 14-23.

LIU Q Y, LIU X F, SUN SH H. Hyperspectral image classification based on shadow enhancement and attention mechanism [J]. Electric Measurement Technology, 2024, 46(8): 14-23.

[11] 刘华咏, 黄聪, 金汉均. 注意力增强的视觉 Transformer 图像检索算法[J]. 电子测量技术, 2024, 46(23): 50-55.

LIU H Y, HUANG C, JIN H J. Attention enhanced visual Transformer image retrieval algorithm[J]. Electric Measurement Technology, 2024, 46(23): 50-55.

- [12] NIU Z, ZHONG G, YU H. A review on the attention mechanism of deep learning [J]. *Neurocomputing*, 2021, 452: 48-62.
- [13] HU J, SHEN L, SUN G. Squeeze-and-excitation networks[C]. *IEEE Conference on Computer Vision and Pattern Recognition*, 2018: 7132-7141.
- [14] GAO Z L, XIE J T, WANG Q L, et al. Global second-order pooling convolutional networks [C]. *IEEE Conference on Computer Vision and Pattern Recognition*, 2019: 3024-3033.
- [15] ZHANG Q L, YANG Y B. SA-Net: Shuffle attention for deep convolutional neural networks [C]. *IEEE International Conference on Acoustics, Speech and Signal Processing*, 2021: 2235-2239.
- [16] HUANG Z L, WANG X G, HUANG L CH, et al. CCNet: Criss-cross attention for semantic segmentation [C]. *IEEE/CVF International Conference on Computer Vision*, 2019: 603-612.
- [17] 陈广秋, 尹文卿, 温奇璋, 等. 基于双重注意力机制生成对抗网络的偏振图像融合[J]. *电子测量与仪器学报*, 2024, 38(4): 140-150.
- CHEN G Q, YIN W Q, WEN Q ZH, et al. Polarization image fusion based on dual attention mechanism for generating adversarial networks[J]. *Journal of Electronic Measurement and Instrumentation*, 2024, 38(4): 140-150.
- [18] WOO S, PARK J, LEE J Y, et al. CBAM: Convolutional block attention module[C]. *European Conference on Computer Vision*, 2018: 3-19.
- [19] BHALLA K, KOUNDAL D, SHARMA B, et al. A fuzzy convolutional neural network for enhancing multi-focus image fusion [J]. *Journal of Visual Communication and Image Representation*, 2022, 84: 103485.
- [20] LIU Y, CHEN X, PENG H, et al. Multi-focus image fusion with a deep convolutional neural network[J]. *Information Fusion*, 2017, 36: 191-207.
- [21] DIAMANTIS D E, IAKOVIDIS D K. Fuzzy pooling[J]. *IEEE Transactions on Fuzzy Systems*, 2021, 29(11): 3481-3488.
- [22] QIN B, NOJIMA Y, ISHIBUCHI H, et al. Realizing deep high-order TSK fuzzy classifier by ensembling interpretable zero-order TSK fuzzy subclassifiers[J]. *IEEE Transactions on Fuzzy Systems*, 2021, 29(11): 3441-3455.
- [23] YANG J CH, SHI R, NI B B. MedMNIST classification decathlon: A lightweight autumn benchmark for medical image analysis [C]. *International Symposium on Biomedical Imaging*, 2021: 191-195.
- [24] NURJANNAH A F, KURNIASARI A S D, SARI Z, et al. Pneumonia image classification using CNN with max pooling and average pooling [J]. *Journal of Information Technology and Systems Engineering*, 2022, 6(2): 330-338.
- [25] WANG Y F, ISHIBUCHI H, PEDRYCZ W, et al. Convolutional fuzzy neural networks with random weights for image classification [J]. *IEEE Transactions on Emerging Topics in Computational Intelligence*, 2024, 8(5): 3279-3293.
- [26] HASAN M M, HOSSAIN M M, RAHMAN M M, et al. FP-CNN: Fuzzy pooling-based convolutional neural network for lung ultrasound image classification with explainable AI[J]. *Computers in Biology and Medicine*, 2023, 165: 107407.
- [27] GU X W, ANGELOV P P. Multiclass fuzzily weighted adaptive-boosting-based self-organizing fuzzy inference ensemble systems for classification[J]. *IEEE Transactions on Fuzzy Systems*, 2021, 30(9): 3722-3735.
- [28] GU X W, ANGELOV P, HAN J G, et al. Multilayer evolving fuzzy neural networks[J]. *IEEE Transactions on Fuzzy Systems*, 2023, 31(12): 4158-4169.
- [29] RANA M S, KABIR M H, SOBUR A. Comparison of the error rates of MNIST datasets using different type of machine learning model[J]. *North American Academic Research*, 2023, 6(5): 173-181.
- [30] KADAM S S, ADAMUTHE A C, PATIL A B. CNN model for image classification on MNIST and Fashion-MNIST dataset [J]. *Journal of Scientific Research*, 2020, 64(2): 374-384.

作者简介

顾苏杭,副教授,主要研究方向为模糊系统、深度学习、数据挖掘等。

E-mail:gusuhang09@163.com

王冶,硕士研究生,主要研究方向为智能系统、图像分类、模糊聚类。

E-mail:wangye14195@163.com

张远鹏,教授,主要研究方向为模式识别、数据挖掘、模糊系统等。

E-mail:y. p. zhang@ieee.org

焦竹青(通信作者),教授,主要研究方向为复杂网络计算、大脑信息处理等。

E-mail:jzp@cczu.edu.cn