QRMatch：用于二维码分类的双域对比与并行自适应增强的半监督框架

摘要：二维码图像在消费电子领域广泛存在，极大地提升智能终端在支付、识别与交互等场景的效率与体验。然而，大量的二维码图像受到环境影响导致扫描设备难以识别各种类型的低质量二维码图像。尽管现有的图像复原方法已取得了有希望的结果，但这些方法不能泛化到所有类别的低质量图像，仍依赖于二维码图像分类后的针对性复原。本文提出了一种识别低质量二维码图像类型的半监督分类网络QRMatch。具体来说，我们观察到不同低质量二维码图像的频谱特征差异明显，有利于学习正、负样本间差异性表征。鉴于此，我们提出了一种双域对比学习方法，通过傅里叶变换将低质量二维码图像切换至频域以挖掘样本间的互补信息；其次，我们引入了一种基于样本历史损失的自适应图像增强模块，重新激活无效的训练样本。此外，为了充分发挥双域对比学习和自适应图像增强模块的优势，我们设计了一种双路径增强数据集流训练框架，缓解了多个优化目标直接相互干扰造成的影响。在二维码数据集上的定量和定性实验结果表明，所提出的方法在三种标签样本比例下优于以往的半监督方法。相较于全监督方法在二维码分类准确率上提高了3.48%，F1分数提升了3.27%。

关键词：二维码分类；半监督学习；傅里叶变换；对比学习；自适应增强

Introduction

二维码（QR Code）凭借其高效的信息存储与传递能力，已成为现代社会中不可或缺的技术，广泛应用于移动支付、物流追踪、身份验证等多个领域[1][2]。随着二维码的普及，真实世界中的二维码图像由于环境因素的影响受到不同程度的劣化导致难以识别。如图1 (a)所示，二维码可能出现模糊、污损、光照不均以及畸变等多种劣化情形。虽然现有的图像复原算法已经能够针对特定类型的劣化实现有效复原，但这些方法不能泛化到所有类别的低质量图像，仍依赖于二维码图像分类后的针对性复原。面对这些挑战，准确分类不同类型的二维码变得尤为重要，它间接影响解码策略和信息提取的精确度，从而提升整体解码成功率和用户体验[4]。

传统的二维码识别和分类方法主要依赖于手工设计的特征提取和规则判断[5,6]。这些方法虽然在理想条件下表现良好，但面对复杂环境中的多样化二维码类型时，其泛化能力和鲁棒性显著下降[7]。近年来，深度学习方法在图像分类领域取得了突破性进展，基于CNN和ViT的全监督学习方法被引入二维码分类任务，显著提升了分类精度。例如，Sarkhi等人[8]提出轻量级深度学习模型识别基于二维码的网络攻击，展示了CNN在二维码异常分类中的有效性。Hu等人[9]设计融合混合注意力机制的轻量级视觉Transformer模型，提高了二维码缺陷分类的准确性和鲁棒性。然而，这些全监督方法依赖于大量的标注数据训练。而在实际应用中，对于低质量二维码的采集和人工标注费时费力[10]，尤其是低光照环境的二维码图像。



图1 随着二维码在真实场景中的广泛应用，分类模型必须对不同类型的低质量劣化二维码图像保持稳健的分类性能，以便应用相应的复原策略。(a) 生活中随处可见的低质量二维码。(b) 常见的二维码扫描设备。(c) 设备难以直接识别低质量的二维码图像。

半监督学习作为一种能够同时利用有标签数据和无标签数据的方法，为解决标注数据有限的问题提供了可行途径[11]。现有的半监督学习方法主要包括伪标签法[12]、一致性正则化[13]以及基于图的方法[14]等。虽然这些方法在多个视觉任务中取得了成功，但直接应用于二维码分类时仍存在局限性：一方面，这些方法通常假设所有样本对模型学习的贡献相同，而二维码图像在实际应用中往往包含难易程度不同的样本，导致模型产生训练偏置；另一方面，不同类型的低质量二维码图像所包含的高频信息差异明显，这些结构特征对于分类至关重要，现有方法未能充分考虑到这些特性。近年来，contrastive learning在视觉表示学习领域展现出巨大潜力，contrastive learning通过挖掘正负样本对实现差异性判别，有效提升了模型的无监督特征表示能力。然而，这些contrastive learning方法(e.g., MOCO[16])主要聚焦于特征域的对比，未能充分利用图像在不同表示空间中的互补信息[17]。如图 2 所示，不同类型低质量二维码图像的频谱特征存在显著差异，这些差异对提升分类性能具有潜在价值，值得深入研究。此外，现有半监督学习方法多采用统一的数据增强策略，忽略了样本难度差异。例如，某些“简单样本”在经过标准数据增强后仍易被网络分类，对模型优化的贡献有限[18]。

图片包含 形状

AI 生成的内容可能不正确。

图2 QR Code V2数据集中4种不同类别二维码的频率域特征，呈现出类型间显著的频谱差异。

针对上述挑战，本文提出了一种基于频域对比学习和自适应图像增强的半监督二维码图像分类方法。具体来说，针对标注数据有限、难以充分挖掘不同表示空间特征的问题，我们提出了一种双域对比学习方法，融合了特征域的MOCO与频率域的对比学习，通过傅里叶变换引入频率域表征，与空间域特征互补，提升模型对二维码结构模式和高频特征的捕捉能力。然后，为了缓解样本难度差异导致的模型学习偏置问题，我们引入了基于历史损失信息的自适应图像增强模块，通过对简单样本实施多样化的强增强策略，提高网络对复杂环境二维码图像的泛化能力。最后，针对不同学习策略间的干扰问题，我们设计了一种双路径增强数据流架构，使双域对比学习与自适应增强模块并行工作且互不干扰，充分发挥了各自优势。

我们在公开的多类型二维码数据集上进行了广泛实验，该数据集包含阴影、弯曲、损坏等多种低质量类型的二维码图像。定性与定量实验结果表明，与现有半监督学习方法相比，所提出的方法在二维码图像识别能力显著提升，分类准确率达到77.39%，F1分数达到78.70%，尤其在标注数据比例低至10%的情况下，相较于对比方法仍展现出一定的性能优势。

Related work

2.1 半监督分类

半监督学习旨在利用少量标注数据和大量未标注数据共同训练模型，在标注资源有限的场景下具有重要价值。近年来，半监督分类方法取得了显著进展。其中，伪标签法[12]作为一种简单而有效的方法，通过利用模型为未标注数据生成伪标签并将其纳入训练过程，从而扩充标注数据。但这类方法高度依赖伪标签的置信度，容易受到噪声标签的影响。另一条重要研究路线是基于一致性正则化的方法，其核心思想是鼓励模型对同一样本在不同数据增强下的预测保持一致。典型代表如Mean Teacher[21]通过构建教师-学生模型并采用EMA更新策略，提升了伪标签的稳定性。UDA[22]则通过显式最小化不同增强视图间的预测差异，有效增强模型的泛化能力。此外，结合伪标签与一致性正则化的方法也展现出优越性能。例如，MixMatch[19]整合了熵最小化、一致性正则化和Mixup数据混合等多种半监督学习技术，显著提升了模型性能。而FixMatch[20]则简化了半监督学习框架，用弱增强视图生成高置信度伪标签并通过强增强视图训练，实现简洁高效的半监督学习。

尽管这些方法在自然图像分类任务中表现优异，但它们依赖于类别间显著的视觉差异（如猫与狗）提取判别特征。二维码图像分类更接近于细粒度分类，各类二维码本质相同，仅在结构与频率特征上存在细微差异。此外，实际场景中的二维码图像往往因模糊、污损、光照不均而质量下降，使传统半监督方法难以在标注有限的情况下准确区分。因此，本文提出融合双域对比学习与自适应增强策略的方法，以更有效地学习不同低质量类型的二维码图像的判别特征。

2.2 对比学习

对比学习近年来在表示学习领域取得了突破性进展。Chen等[23]提出的SimCLR通过最大化同一图像不同增强视图之间的相似性，同时最小化不同图像增强视图之间的相似性，学习丰富的视觉表示。He等[16]提出的MoCo引入动量编码器和队列机制，扩大负样本范围，提高对比学习效果。MoCo v2[24]和MoCo v3[25]通过引入额外的增强策略和改进的架构设计，进一步提升了表示学习能力。对比学习也在半监督场景中展现出显著的应用潜力。例如，SimCLRv2[26]将对比预训练与半监督微调相结合，显著降低了对标注数据的需求。Yang等[27] 提出Class-aware Contrastive Semi-Supervised Learning（CCSSL），通过结合伪标签与类别感知的对比损失，增强模型对同类样本的聚合能力，在复杂分布下提升了判别性能。

然而，现有对比学习方法主要在特征域进行特征提取与对比，忽略了频域信息的潜在价值。对于不同类型的低质量二维码图像，其成像特征在频域中可能具有显著的区分性。本文提出的双域对比学习方法通过同时在特征域和频率域进行对比学习，能够更全面地捕捉不同成像条件下的图像特征，从而提升分类性能。

2.3 数据增强

数据增强是深度学习中提高模型泛化能力的关键技术。传统数据增强方法包括随机裁剪、翻转、旋转和颜色抖动等[28]。近年来，更复杂的增强策略如AutoAugment[29]和RandAugment[30]通过搜索或随机组合多种基本增强操作，实现了更有效的数据增强。在此基础上，自适应数据增强方法逐渐兴起，它们根据样本特性或训练过程动态调整增强强度与策略。例如，Ho等[31]提出的Population Based Augmentation通过进化算法搜索最优增强策略。Ahn等[32]提出基于类别样本数量自适应调整增强策略，对长尾分布中的少数类施加更强增强。相反，Lin等[33]发现对"简单样本"实施强增强可能更有利于模型性能，这与传统认知形成了有趣的对比。

在半监督学习中，数据增强策略的选择对模型性能有显著影响。FixMatch[20]使用弱增强生成伪标签，强增强用于训练，展示了不同增强强度的组合价值。然而，大多数现有方法仍采用固定的增强策略，缺乏对样本动态特性的建模。已有的自适应增强策略主要依据类别或样本复杂度进行调整，较少考虑样本在训练过程中的学习动态（如历史损失变化）。为此，本文引入一种基于样本历史损失的信息驱动增强策略，识别"简单样本"并对其施加多样化的强增强策略，提升模型对挑战性样本的学习能力。结合所设计的双路径增强数据流架构，该策略可与双域对比学习协同工作，共同增强二维码分类的性能。

Method

3.1 网络架构

QRMatch的网络架构如图3所示。我们定义有标签数据和无标签数据为：



其中，代表有标签的二维码样本，是这些样本的类别标签，代表无标签的二维码样本。对于每个，其伪标签将根据模型对该数据的预测结果来自动生成。

图示

AI 生成的内容可能不正确。

图3 Overview of QRMatch。监督学习分支利用有标签数据训练模型，优化分类性能，并为无监督学习中的无标签数据提供参考依据。无监督学习分支采用双路径增强数据流架构，将数据流分为两个并行路径，确保双域对比学习模块和样本自适应增强模块能够协同工作而不相互干扰。具体来说，弱增强样本和强增强样本被输入到双域对比学习模块，分别在特征域和频率域进行对比学习，以提升模型的鲁棒性和特征的区分能力；与此同时，弱增强样本和强增强样本的副本被用于计算一致性损失，样本自适应增强模块基于这些副本的历史一致性损失识别出“简单样本”，并对这些样本进一步应用强增强，生成新的样本，从而增强模型对挑战性样本的学习能力。

对于标注样本，对其进行随机裁剪、翻转等弱数据增强后，将增强样本输入神经网络获得预测输出，并通过与真实标签的交叉熵损失更新模型参数。对于未标注样本，对其应用强增强和弱增强处理，随后增强样本进入编码器提取特征，模型输出预测结果并计算一致性损失。为充分利用无标签数据并提升模型对图像细节的敏感性，我们在强、弱增强数据之间引入了双域对比学习：特征域对比学习聚焦于语义特征的一致性，频率域对比则强调细节特征的捕捉，二者结合使模型能够兼顾语义和细节，形成更全面的特征表示。同时我们设计了样本自适应增强模块，通过监控强增强数据副本的历史一致性损失来识别简单样本。当低于某一阈值时，表明该样本对模型学习的贡献较小，我们会对强增强样本的副本施加进一步的强增强，以提升其训练价值。此外，我们还设计了创新的双路径增强数据流架构，为确保双域对比学习与样本自适应增强模块协同工作，我们提出了创新的双路径增强数据流架构。该架构将数据流分为两个并行路径，避免了模块间的相互干扰，使各组件充分发挥作用。

总之，QRMatch能够有效实现半监督二维码图像分类。之后，我们将在3.2节介绍双路径增强数据流架构，并在3.3和3.4节介绍双域对比学习模块与样本自适应增强模块。

3.2 双路径增强数据流架构

在半监督分类网络中，为了充分利用无标签数据并提升模型性能，我们引入了两个创新模块：双域对比学习模块和样本自适应增强模块。然而，这两个模块均依赖于对强增强数据的操作——前者通过强增强数据与弱增强数据之间的对比学习增强表示能力，后者通过对强增强数据的再增强提升模型对简单样本的贡献度——这可能导致训练过程中的学习不稳定性。为此，我们提出了一种新颖的“双路径增强数据流架构”，旨在通过精心设计的数据流路径，协调对比学习和一致性正则化之间的关系，同时动态优化样本增强策略，从而显著提升模型的稳定性和分类性能。

“双路径增强数据流架构”的核心在于将强增强数据流分为两个独立的并行路径，每条路径专注于特定的学习任务。这种分离设计避免了在单一数据流上叠加多个增强操作可能引发的相互干扰，同时为对比学习和一致性正则化的独立优化提供了灵活性。此外，通过结合样本自适应增强策略，该架构能够动态调整模型对数据难度的感知，从而更高效地利用无标签数据。

给定一个无标签样本，我们首先通过数据增强生成其弱增强版本和强增强版本。随后，我们为强增强数据复制一个副本：​，以用于不同的学习路径。在对比学习路径中，输入是原始强增强数据和弱增强数据，在这一路径中，我们应用双域对比学习模块，分别在特征域和频率域计算和之间的对比损失，通过这种跨域对比学习的方法，强化模型对弱增强和强增强数据之间一致性表示的学习。而在一致性正则化路径中，输入是复制的强增强数据和弱增强数据，我们在这两个数据的网络预测之间应用一致性正则化损失，然后利用样本自适应增强模块识别出对模型学习贡献较小的简单样本，并对这些样本的应用额外的再增强操作生成，并用替换原始的，确保模型自适应聚焦挑战样本，提高泛化能力并优化训练稳定性。

3.3 双域对比学习模块

本研究提出的双域对比学习模块是一种新颖的表示学习方法，如图4所示，该方法同时在特征域和频率域进行对比学习，从而获取二维码图像更为鲁棒和丰富的特征表示。该框架基于动量对比学习的基础架构，并创新性地引入频率对比正则化，形成了一种双域协同的表示学习范式。

图示

AI 生成的内容可能不正确。

图4 双域对比学习模块

特征域分支采用了动量对比学习框架，包含查询编码器和键编码器两个网络。两个编码器共享相同的网络结构，但参数更新方式不同：查询编码器通过梯度下降直接更新，键编码器则通过动量更新机制进行参数更新：



其中，和分别表示键编码器和查询编码器的参数，为动量系数，控制参数更新速率。

在训练过程中，查询图像和键图像（两者为同一图像的不同增强版本）分别通过各自的编码器，生成标准化的特征向量和。系统维护一个大小为的负样本队列，通过以下对比损失函数优化网络：



其中，是温度系数，为队列中的负样本特征。

在频率域分支中，我们引入频率对比正则化(FCR)机制，以在傅里叶频率域中构建对比学习任务。该机制包含两个核心组件：频率域负样本池维护策略和频域对比计算。首先，为保障频域对比学习的有效性，我们设计了专门的负样本池维护机制：模型维护一个包含最新负样本原始图像的队列，该队列与MoCo原有的特征队列并行运作，然后我们采用先进先出(FIFO)的策略维护图像队列，确保始终保留最新的K个样本图像：



其中，表示在第步时，队列中存储的负样本原始图像集合，表示在第步新产生的负样本图像， 表示将新的负样本加入队列末尾，表示仅保留末尾最近的个元素（FIFIO策略），当队列长度超过时，丢弃最早的元素，使得队列最终大小保持在以内。在每次前向传播时，模型从队列中提取最近的K个负样本图像用于频率域对比学习：



这里，表示第t布前向传播中用到的负样本图像集合，这样我们就从中获取了最新个元素，确保频率域对比学习始终使用最新的负样本。

然后，我们进行频率域对比计算。我们对锚点图像、正样本图像以及从负样本队列中选取的图像执行执行二维快速傅里叶变换(2D-FFT)，获得它们在频率域的表示：



其中，表示傅里叶变换。接着，在频率域中，我们计算锚点与正样本之间的L1距离，以及锚点与多个负样本之间的L1距离：



最后我们构建频率域对比损失函数，使锚点与正样本在频率域的距离相对于与负样本的距离更小：



其中，为批次大小，为每个锚点选择的负样本数量，为数值稳定性常数。

最终，模型通过联合优化特征域对比损失和频率域对比损失来学习二维码的鲁棒表示：



其中，为平衡两个损失项的权重系数。

这种双域对比学习框架通过融合空间结构与频率特征，有效增强模型对噪声、模糊及几何畸变的鲁棒性。其频率域模块通过分析不同场景二维码图像的独特频谱特性，有效区分二维码类型。这种多视角特征理解机制支持在有限标注数据下实现高效半监督学习。

3.4 样本自适应增强模块

样本自适应增强模块（SAA）旨在动态调整增强策略，针对性地处理训练过程中的“简单样本”。这些样本往往经增强后仍能被网络以高置信度正确分类，其损失近乎为零，从而阻碍了模型的进一步学习。SAA通过两个组件来实现对简单样本的优化：样本选择模块和样本增强模块。

如图5上方所示，样本选择模块通过追踪未标记样本的历史一致性损失（Historical Consistency Loss）并借助OTSU算法，在每轮训练中自适应地挑出对模型贡献不大的“简单样本”。由于不同拍摄场景下的二维码图像在外观、学习难度上存在显著差异，全局的OTSU阈值会导致无法精确区分不同场景的学习难度差异。受Flexmatch的启发，我们开发了一种类特异性阈值OTSU算法，为每个类别动态计算独立的阈值。

图片包含 图形用户界面

AI 生成的内容可能不正确。

图5 样本自适应增强模块的流程示意图

首先，我们在训练过程中记录每个样本的损失值，并根据模型对该样本的预测类别进行分组。例如，对于一个无标签样本，模型预测其属于类别（），则将该样本的当前一致性损失加入类别的历史损失集合中，这样我们可以得到所有类别的历史损失集合。然后，我们使用指数移动平均（EMA）在每个训练轮次更新每个样本的历史损失：



其中代表平滑因子，是第t个epoch的样本一致性损失。通过这种方法，SAA能够平滑样本损失的波动，更好地反映出样本对模型训练的长期影响。然后我们对每个类别c的历史损失集合应用OTSU算法，计算该类别的动态阈值。OTSU算法会根据的分布自适应地将样本分为两组（简单和非简单），其核心是最大化类间方差。得到的是类别c专属的阈值，而不是全局统一的阈值。

对于每个样本，获取模型预测类别（其中是弱增强版本的预测概率），然后根据的历史损失和对应类别的阈值判断：若，则标记为简单样本，否则标记为非简单样本。

接下来，我们在样本增强模块中为简单样本应用更复杂的增强策略。对于简单样本，我们采用一种新的增强策略，该策略通过将两个独立的增强版本拼接成一个新的图像，从而增加图像的多样性，公式为：







其中， 和是无标签样本的两个独立增强版本，代表高斯噪声操作，是高斯噪声的方差。代表随即掩码操作，是一个掩码图像，代表哈达玛乘积，代表拼接操作。由于高斯噪声往往用于模拟拍照时的传感器噪点、弱光噪声或图像噪声；随机掩码可以模拟局部遮挡、污渍、划痕、反光亮斑等，这两个增强手段非常贴近真实拍摄中出现的干扰因素，因此我们在样本自适应模块中使用这两种方法作为再增强操作。

这种方法可以大幅度增加增强版本的多样性，从而促使模型更好地从这些样本中学习。最终，SAA通过动态调整每个样本的增强策略，确保简单样本得到充分的利用，避免它们对训练的贡献过低，从而有效提升了模型的性能。

Experiments

4.1 实施细节Implementation Details

在训练阶段，我们将初始学习率设置为=0.03，并实施余弦学习率衰减策略[35]以渐进式降低学习率，从而提升模型训练效果与稳定性。优化过程中，我们采用动量参数为0.9的SGD[36]优化器，通过梯度下降算法在每次迭代中更新模型参数，其中k和K分别代表当前训练步数和总训练步数，设定为k=30和K=15000。为确保评估的公平性，所有方法均使用ResNet-18[37]作为骨干网络架构，并将标注数据的批量大小统一设为16。对于QRMatch，由于训练初期模型尚未收敛，可能对某些样本产生不准确的置信度估计，导致“一致性损失”出现极端值，从而错误地将样本分类为"简单"或"非简单"。因此我们引入“warm-up”阶段，在训练初期100轮内暂不启用样本自适应增强，以避免模型过早地对样本进行刚性划分。此外，我们使用Python 3.10和PyTorch 2.1.1实现所有方法，并在一台配备 12 核 Intel Xeon Silver 4214R 2.4 GHz 处理器（128GB 内存）和四张 NVIDIA GeForce RTX 3090 24G GPU 的计算机上进行训练和测试。

4.2 数据集Dataset

在本研究中，我们使用QR Code V2[38]这个公开的二维码数据集进行实验，该数据集包含529张在多种真实世界场景下采集的QR码图像样本，这些样本涵盖了多种不同的拍摄条件。如图6所示，该数据集包含“Blurred”、“Brightness”、“Bright Spots”、“Close”、“Curved”、“Damaged”、“Glare”、“High Version”、“Monitor”、“Nominal”、“Non-Compliant”、“Pathological”、“Perspective”、“Rotated”**、**“Shadows”这15个类别。由于该数据集中不同类别的图像分辨率并不相同，因此我们在实验过程中统一将图像分辨率调整为256×256，以确保模

QR 代码

AI 生成的内容可能不正确。

图6 QR Code V2数据集中的各类别二维码图像示例

型输入尺寸一致，同时提高训练的稳定性和计算效率。受以往工作的启发[39][19]，我们使用不同数量的标记样本进行了实验。

4.3 实验结果和分析

我们在QR Code V2数据集上进行了广泛的实验，使用三种不同数量的标签（153个标签（约50%）、95个标签（约30%）、32个标签（约10%））来评估各种半监督分类算法的性能，每个类别的标签分配比例与其总规模成正比。对于样本量较小的类别，即使无法达到10%的划分比例，我们也确保至少分配一个有标签样本。我们使用Top-1、Top-5以及F1分数作为评估指标。

各算法的实验结果如表1所示。QRMatch在三种不同的标签设置下均取得了令人满意的结果。值得注意的是，当标签数量从153个减少至95个时，QRMatch的Top-1准确率仅从77.39%下降至73.05%，下降了4.34%，而Top-5准确率下

表1：在QR Code V2数据集上与最先进的半监督分类算法进行比较。Top-1、Top-5准确率(%)以及F1分数(%)作为评价指标

形状

AI 生成的内容可能不正确。

降了3.02%，F1分数下降了5.55%。相比之下，在153个标签下性能仅次于QRMatch的Softmatch，在95个标签设置下的Top-1准确率下降了5.1%，Top-5准确率和F1分数分别下降了3.94%和6.23%。受标签样本数量的影响，Fixmatch、Flexmatch和Mixmatch的Top1准确率分别下降了6.93%、5.28%和5.21%，F1分数分别下降了9.67%、7.13%和6.69%。这种性能下降源于标签数量减少，导致有标签样本的多样性和代表性下降，进而削弱了半监督学习算法学习鲁棒特征的能力。

作为新颖的半监督分类算法，Freematch和Soc在153个有标签数据的设置下分别取得了73.04%和74.74%的Top-1准确率，以及73.91%和75.62%的F1分数。当标签数量减少至95个时，Freematch受到的影响略小于Soc：Freematch的Top-1准确率下降了4.46%，F1分数下降了5.72%，而Soc的Top-1准确率下降了5.42%，F1分数下降了7.46%。进一步地，当有标签样本数量进一步缩减至28个时，Freematch的Top-1准确率为53.94%，F1分数为53.02%，而Soc的Top-1准确率为53.5%，F1分数为52.32%。这表明在低标签样本数量的情况下，Soc方法的性能衰减比Freematch更为显著。Soc方法主要依赖类别转换跟踪和置信度感知聚类生成伪标签，当样本数量不足时，类别转换跟踪估计不准确，聚类结果不可靠，导致性能下降。相比之下，Freematch依赖置信度阈值和一致性正则化，不涉及复杂的类相似性估计，因此对标签样本数量减少表现出更强的鲁棒性。

由于验证集与训练数据在统计特征上保持一致，因此我们选择在32个有标签数据的设置下，对训练期间三种方法预测验证集样本的准确率进行统计分析。如图7所示，在训练初期，Fixmatch、Softmatch和QRMatch预测验证集上的准确率均处于较低水平。随着迭代次数的增加，QRMatch和Softmatch的准确率呈现波动上升趋势，而Fixmatch受限于其严格的伪标签阈值筛选策略，在少样本条件下模型的泛化能力明显受限。

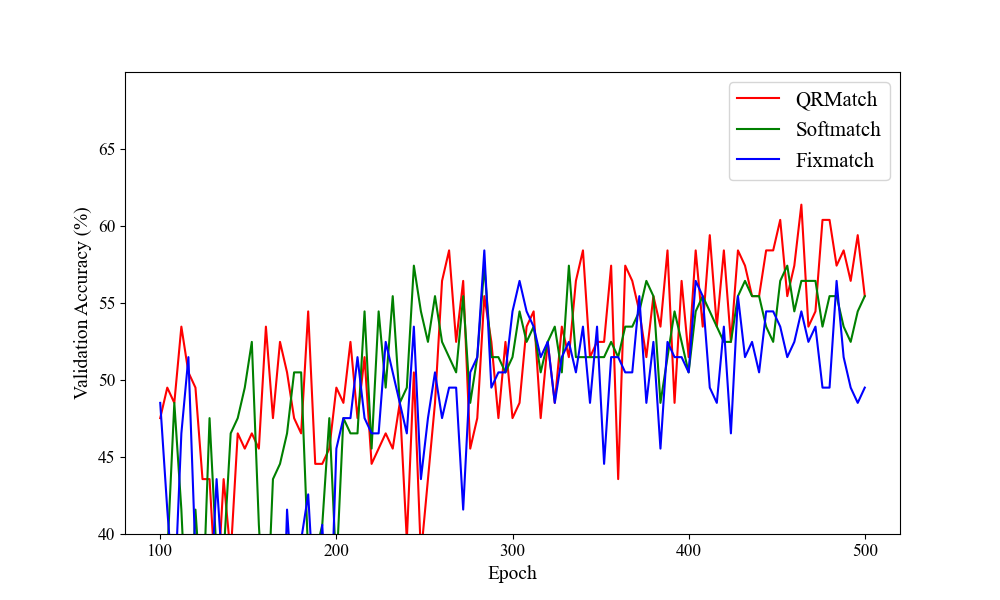


图7：当只有32个标签用于训练时，Fixmatch、Softmatch和QRMatch在验证集上的准确率

此外，我们在图8中提供了在153个有标签数据的设置下，各分类算法的特征空间可视化结果。显然，QRMatch相较于其他对比算法在二维码多类别识别任务中展现出更优越的特征分离能力。尽管15类样本的复杂性导致UMAP投影存在部分重叠区域，但结合混淆矩阵分析可知，QRMatch在关键挑战性类别上实现了更紧凑的类内分布和更显著的类间边界。具体而言，Brightness（第2类）、High Version（第8类）和Pathological（第12类）等结构特征明显的类别，所有方法均能保持较高的识别率，其特征在低维空间中呈现高度聚集性。然而，对于受光照干扰的Glare（第7类）和Shadows（第15类），QRMatch的准确率显著优于Fixmatch（含跨类干扰）和Flexmatch（存在严重误判）。特别值得注意的是，在Blurred（第1类）和Nominal（第10类）这类易混淆类别中，QRMatch展现出优于Softmatch（误判分散至多个类别）和Freematch（存在明显的跨类干扰）的识别能力，表明其能有效捕捉模糊退化与标准形态间的细微差异。尽管Mixmatch在Damaged（第6类）取得一定识别率，但其对Curved（第5类）的样本频繁误分为Nominal，而QRMatch通过较高的准确率验证了对物理形变特征的鲁棒编码能力。

散点图

AI 生成的内容可能不正确。

图8：UMAP (sizes=25, n\_neighbors=10, min\_dist=1.0) feature visualization on QR Code V2 dataset.

4.4 消融实验

如表2所示，我们在QR Code V2数据集的153个标签设置下进行消融实验，以验证QRMatch各组件的有效性。Index 1代表在baseline上的性能指标，Index 2代表在Index 1的基础上引入常规的动量对比学习，Index 3代表在在Index 2的基础上引入频率域对比学习分支，构成双域对比学习模块，Index 4代表在Index 1的基础上引入样本自适应增强，Index 5代表在Index 3的基础上引入样本自适应增强，Index 6代表在Index 5的基础上引入我们提出的双路径增强数据流架构。实验结果表明，双域对比学习模块和样本自适应增强模块均能显著提升

表2 ：QRMatch的组件消融研究和监督学习下的对比实验

形状

AI 生成的内容可能不正确。

二维码半监督分类的准确性。Index 2和Index 3的实验结果表明，增加频率域对比学习分支对于二维码半监督分类是有效的。值得注意的是，简单地组合双域对比学习(DDCL)和样本自适应增强(SAA)并不能提高分类性能，反而会严重损害模型表现。正如我们此前论述，无标签样本同时接受DDCL和SAA处理会导致过度扰动，不利于模型学习。而我们提出的双路径增强数据流架构（Index 6）有效地解决了上述问题，这种架构巧妙地平衡了数据增强的力度和模型的泛化能力，体现了我们在方法设计上的创新性。

此外，Index 7和Index 8分别表示使用153个标签和使用所有训练样本标签的Resnet-18消融实验。实验结果显示，在半监督学习中，QRMatch在三个主要指标上全面超越了全监督的Resnet-18。QRMatch通过双域对比学习模块(DDCL)的跨域特征学习、SAA增强无标签数据贡献和DP-EDFA的协同优化，有效挖掘了数据的内在结构，提升了半监督分类的性能。

Conclusion

本文提出了一种新型的半监督学习框架 QRMatch，专注于二维码分类任务中的准确性和鲁棒性提升。我们设计了 双域对比学习（DDCL），结合MOCO机制的特征域对比和傅里叶变换的频率域对比，以充分挖掘二维码图像的多层次信息。此外，我们引入了样本自适应增强（SAA），基于样本的历史损失动态调整增强策略，从而提升对不同难度样本的学习能力。为了确保不同模块协同工作，我们提出双路径增强数据流架构（DP-EDFA），有效减少了增强数据流之间的相互干扰。未来，我们将探索QRMatch在更大规模和更复杂场景下的泛化能力，并结合自监督预训练和Transformer结构进一步提升二维码分类的鲁棒性和适应性。

References：

Introduction部分

[1] Eren B A. QR code m-payment from a customer experience perspective[J]. Journal of Financial Services Marketing, 2022: 1.

[2] Scanzio S, Rosani M, Scamuzzi M, et al. QR codes: From a Survey of the State-of-the-Art to Executable eQR Codes for the Internet of Things[J]. IEEE Internet of Things Journal, 2024.

~~[3] Huo L, Zhu J, Singh P K, et al. Research on QR image code recognition system based on artificial intelligence algorithm[J]. Journal of Intelligent Systems, 2021, 30(1): 855-867.~~

[4] Munoz-Mejias D, Gonzalez-Diaz I, Diaz-de-Maria F. A low-complexity pre-processing system for restoring low-quality QR code images[J]. IEEE Transactions on Consumer Electronics, 2011, 57(3): 1320-1328.

[5] Liao Z, Huang T, Wang R, et al. A method of image analysis for QR code recognition[C]//2010 International Conference on Intelligent Computing and Integrated Systems. IEEE, 2010: 250-253.

[6] Liu Y, Liu M. Automatic recognition algorithm of quick response code based on embedded system[C]//Sixth International Conference on Intelligent Systems Design and Applications. IEEE, 2006, 2: 783-788.

[7] Yang Z, Xu H, Deng J, et al. Robust and fast decoding of high-capacity color QR codes for mobile applications[J]. IEEE Transactions on Image Processing, 2018, 27(12): 6093-6108.

换了，换成近年的文献，且就是做二维码分类的[8] Sarkhi M, Mishra S. Detection of QR code-based cyberattacks using a lightweight Deep Learning model[J]. Engineering, Technology & Applied Science Research, 2024, 14(4): 15209-15216.

换了，换成近年的文献，且是做二维码分类[9] Hu D, Zhao L, Ren Y, et al. A lightweight vision transformer with embedded hybrid attention for quick response code defect classification[J]. Engineering Applications of Artificial Intelligence, 2025, 150: 110575.

换成近年的了[10] Tsai M J, Lee Y C, Chen T M. Implementing Deep Convolutional Neural Networks for QR Code-Based Printed Source Identification[J]. Algorithms, 2023, 16(3): 160.

[11] Yang X, Song Z, King I, et al. A survey on deep semi-supervised learning[J]. IEEE Transactions on Knowledge and Data Engineering, 2022, 35(9): 8934-8954.

[12] Cascante-Bonilla P, Tan F, Qi Y, et al. Curriculum labeling: Revisiting pseudo-labeling for semi-supervised learning[C]//Proceedings of the AAAI Conference on Artificial Intelligence. 2021, 35(8): 6912-6920.

[13] Chen K, Wang S. Semi-supervised learning via regularized boosting working on multiple semi-supervised assumptions[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2010, 33(1): 129-143.

[14] Deng J, Yu J G. A simple graph-based semi-supervised learning approach for imbalanced classification[J]. Pattern Recognition, 2021, 118: 108026.

~~[15] Khare S K, Bajaj V. Time–frequency representation and convolutional neural network-based emotion recognition[J]. IEEE Transactions on Neural Networks and Learning Systems, 2020, 32(7): 2901-2909.~~

[16] He K, Fan H, Wu Y, et al. Momentum contrast for unsupervised visual representation learning[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2020: 9729-9738.

[17] Lin Y, Gou Y, Liu X, et al. Dual contrastive prediction for incomplete multi-view representation learning[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2022, 45(4): 4447-4461.

[18] Gui G, Zhao Z, Qi L, et al. Enhancing sample utilization through sample adaptive augmentation in semi-supervised learning[C]//Proceedings of the IEEE/CVF International Conference on Computer Vision. 2023: 15880-15889.

相关工作/半监督分类

[19] Berthelot D, Carlini N, Goodfellow I, et al. Mixmatch: A holistic approach to semi-supervised learning[J]. Advances in Neural Information Processing Systems, 2019, 32.

[20] Sohn K, Berthelot D, Carlini N, et al. Fixmatch: Simplifying semi-supervised learning with consistency and confidence[J]. Advances in Neural Information Processing Systems, 2020, 33: 596-608.

[21] Tarvainen A, Valpola H. Mean teachers are better role models: Weight-averaged consistency targets improve semi-supervised deep learning results[J]. Advances in Neural Information Processing Systems, 2017, 30.

[22] Xie Q, Dai Z, Hovy E, et al. Unsupervised data augmentation for consistency training[J]. Advances in Neural Information Processing Systems, 2020, 33: 6256-6268.

相关工作/对比学习

[23] Chen T, Kornblith S, Norouzi M, et al. A simple framework for contrastive learning of visual representations[C]//International Conference on Machine Learning. PmLR, 2020: 1597-1607.

[24] Chen X, Fan H, Girshick R, et al. Improved baselines with momentum contrastive learning[J]. arXiv preprint arXiv:2003.04297, 2020.

[25] Chen X, Xie S, He K. An empirical study of training self-supervised vision transformers[C]//Proceedings of the IEEE/CVF International Conference on Computer Vision. 2021: 9640-9649.

2020年[26] Chen T, Kornblith S, Swersky K, et al. Big self-supervised models are strong semi-supervised learners[J]. Advances in Neural Information Processing Systems, 2020, 33: 22243-22255.

2022年[27] Yang F, Wu K, Zhang S, et al. Class-aware contrastive semi-supervised learning[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2022: 14421-14430.

相关工作/数据增强

[28] Krizhevsky A, Sutskever I, Hinton G E. Imagenet classification with deep convolutional neural networks[J]. Advances in Neural Information Processing Systems, 2012, 25.

[29] Cubuk E D, Zoph B, Mane D, et al. Autoaugment: Learning augmentation policies from data[J]. arXiv preprint arXiv:1805.09501, 2018.

[30] Cubuk E D, Zoph B, Shlens J, et al. Randaugment: Practical automated data augmentation with a reduced search space[C]//Proceedings of the IEEE/CVF conference on computer vision and pattern recognition workshops. 2020: 702-703.

[31] Ho D, Liang E, Chen X, et al. Population based augmentation: Efficient learning of augmentation policy schedules[C]//International conference on machine learning. PMLR, 2019: 2731-2741.

[32] Ahn S, Ko J, Yun S Y. Cuda: Curriculum of data augmentation for long-tailed recognition[J]. arXiv preprint arXiv:2302.05499, 2023.

[33] Lin C H, Kaushik C, Dyer E L, et al. The good, the bad and the ugly sides of data augmentation: An implicit spectral regularization perspective[J]. Journal of Machine Learning Research, 2024, 25(91): 1-85.

方法/3.4

[34] Otsu N. A threshold selection method from gray-level histograms[J]. Automatica, 1975, 11(285-296): 23-27.

实验部分/4.1

[35] Loshchilov I, Hutter F. Sgdr: Stochastic gradient descent with warm restarts[J]. arXiv preprint arXiv:1608.03983, 2016.

[36] LeCun Y, Bottou L, Bengio Y, et al. Gradient-based learning applied to document recognition[J]. Proceedings of the IEEE, 1998, 86(11): 2278-2324.

[37] He K, Zhang X, Ren S, et al. Deep residual learning for image recognition[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2016: 770-778.

实验部分/4.2

[38] @misc{QRBenchmark,

title = {Study of QR Code Scanning Performance in Different Environments. V3},

howpublished = {\url{https://boofcv.org/index.php?title=Performance:QrCode}},

author = {Peter Abeles},

originalyear = {03.19.2019}

}

[39] Li Z, Lin Q, Fan H, et al. SIAVC: Semi-Supervised Framework for Industrial Accident Video Classification[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2024.