

## 摘要

燃气管道作为城市基础设施的重要组成部分，其焊接的安全性、稳定性和可靠性日益受到重视。然而对于焊接缺陷的检测却面临许多挑战：焊接接头的缺陷一般出现在管道的内侧，必须要通过超声波或 X 射线的方式进行探测，且需要具有专业知识与技能的工作人员对图像进行人工识别分类并分级，不仅效率低下还十分容易出现误判漏判。本文希望能够搭建一个系统，其不仅能基于超声相控阵系统完成缺陷的探测，还能够适应数据分布的不断变化，持续地优化已有模型。为了实现以上所提到的功能，我们希望对广泛应用于其他深度学习领域的持续学习方法在焊接缺陷检测的具体任务中进行评估和优化使其能够适应燃气管道焊接缺陷检测的场景。

**关键词：**持续学习，缺陷检测，机器视觉，深度学习

## Abstract

As a vital component of urban infrastructure, gas pipelines' welding safety, stability, and reliability are gaining increasing attention. However, the detection of welding defects poses numerous challenges: these defects typically occur on the inner side of pipe joints, necessitating inspection through ultrasonic or X-ray methods. Furthermore, expert personnel with specialized knowledge and skills must manually identify, classify, and grade defect images, a process that is not only inefficient but also prone to misjudgments and omissions. This paper aims to establish a system capable not only of detecting defects using phased array ultrasonic technology but also of adapting to the continually evolving data distribution, thereby perpetually refining existing models. To attain the functionalities outlined above, we aspire to evaluate and optimize the application of continuous learning methodologies, commonly employed in diverse deep learning domains, to the specific task of welding defect detection. Our objective is to tailor these methodologies to the context of gas pipeline welding defect inspection, ensuring enhanced efficacy and precision.

**Keywords:** Continual Learning, Defect Detection, Machine Vision, Deep Learning

## 目录

摘要.....	I
Abstract.....	II
1 研究背景.....	1
2 文献综述.....	2
1.1 二级标题（黑体四号，左对齐，段前距为12磅）.....	11
1.1.1 三级标题（黑体小四号，左对齐，段前距为12磅）.....	11
1.2 二级标题（黑体四号，左对齐，段前距为12磅）.....	11
1.2.1 三级标题（黑体小四号，左对齐，段前距为12磅）.....	11
2 一级标题（换页并首行居中，黑体小二号，段前距、段后距分别为12磅） .....	12
2.1 二级标题（黑体四号，左对齐，段前距为12磅）.....	12
2.1.1 三级标题（黑体小四号，左对齐，段前距为12磅）.....	12
2.2 二级标题（黑体四号，左对齐，段前距为12磅）.....	12
2.2.1 三级标题（黑体小四号，左对齐，段前距为12磅）.....	12
3 一级标题（换页并首行居中，黑体小二号，段前距、段后距分别为12磅） .....	13
参考文献（换页，黑体小二号居中，段前距、段后距分别为12磅）....	14
致谢（换页，黑体小二号居中，段前距、段后距分别为12磅）.....	14

## 1 研究背景

随着城市化和工业化的快速发展，燃气管道作为城市基础设施的重要组成部分，其安全性、稳定性和可靠性日益受到重视。在燃气管道的施工过程中，焊接是一种常见的连接方式，然而由于各种因素的影响，焊接过程中可能会出现各种缺陷，如未熔合、裂纹、气孔等。这些缺陷可能导致燃气泄漏、管道破裂等安全事故，严重威胁人民生命财产安全。因此，对燃气管道焊接缺陷的检测和预防至关重要。

然而，焊接接头的缺陷一般出现在管道的内侧，无法通过肉眼观测。为了解决这一问题，目前大多使用 X 射线与超声波来进行无损检测。其中 X 射线多用于钢制管道，而超声波多用于 PE 管道的检测。虽然 X 射线或超声相控阵系统能够清晰展示管道内壁的情况，但是仍然需要具有专业知识与技能的工作人员对图像进行人工识别分类并分级，不仅效率低下还十分容易出现误判漏判。

近年来，随着人工智能和机器学习技术的快速发展，机器视觉和持续学习（Continuous Learning）作为重要的机器学习方法，逐渐在各个领域得到应用。在燃气管道焊接缺陷检测领域中，机器视觉技术可以代替人力完成对无损检测图像的识别，提高效率与准确度。然而，燃气管道焊接过程中产生的缺陷数据是动态变化的，随着时间的推移和技术的进步，缺陷的类型和分布可能发生变化，导致原有的机器视觉模型无法适应不断变化的场景。而从头重新训练一个能够适应当前分布的模型不仅会耗费大量时间与算力资源，还可能会使模型对之前缺陷的检测性能发生下降。

持续学习的应用旨在使机器视觉模型能够在不断接收新数据的情况下，持续更新和优化自身的性能，以适应环境的变化。在燃气管道焊接缺陷检测领域，持续学习方法可以在不重新训练整个模型的情况下，仅对新增数据进行学习，从而大大提高学习效率。这对于燃气管道焊接缺陷检测来说非常重要，因为在实际应用中，往往需要对大量的数据进行处理和分析，而持续学习可以在保证准确性的同时，大大提高处理速度，满足实际应用的需求。同时企业可以在不增加额外硬件和软件成本的情况下，实现对燃气管道焊接缺陷检测的不断优化。这不仅可以降低企业的运营成本，还可以减少因管道事故带来的经济损失和社会风险。

本文希望通过使用持续学习技术来扩展机器视觉模型在燃气管道焊接缺陷检测领域的实际应用场景，最后搭建一个能够提供自动化整理在线与离线数据并为视觉模型提供持续优化训练与评估服务的后端平台。为保障城市基础设施的安全稳定运行提供有力支持。

## 2 文献综述

### 2.1 图像缺陷检测

缺陷检测是一个重要的机器学习问题。与大多数现有机器学习方法所基于的静态和封闭系统的假设不同，它研究机器学习模型如何在开放和动态系统环境下处理未知和不确定的信息。在开放环境的假设下，为异常检测开发的学习系统通常期望利用已知（正常数据和模式）的知识来推断未知（与正常不同的异常或新模式）。异常检测方法通常使用可用的正常数据提取、表征和建模模式，然后开发合理的异常检测器来发现新观察到的数据中的新颖或异常模式。当异常检测的目标是图像数据时，就产生了视觉缺陷检测或图像异常检测的任务。

在视觉异常检测中，异常样本或模式可能以多种形式出现，如形状、颜色、纹理或运动模式的异常。这些异常可能由多种因素引起，如摄像头故障、物体异常行为或场景中的罕见事件。因此，视觉异常检测需要能够识别与正常模式显著不同的模式，即使它们可能非常复杂且难以预定义。

为了实现这一点，视觉异常检测方法通常依赖于特征提取和建模技术，以从正常数据中捕获关键信息，并构建能够区分正常和异常模式的模型。这些技术可能包括传统的图像处理技术（如滤波、边缘检测等）、深度学习技术（如卷积神经网络、自编码器等）或统计学习方法（如概率模型、密度估计等）。

此外，视觉异常检测还面临一些挑战，如异常模式的多样性和不可预测性、正常数据和异常数据之间的不平衡、以及缺乏足够的标记数据来训练模型。因此，研究人员一直在努力开发更强大和适应性更强的方法，以在各种场景中实现准确的视觉异常检测。视觉异常检测是一个活跃的研究领域，具有广泛的应用前景，包括安全监控、医疗诊断、工业质量控制等。随着技术的不断进步，我们相信在未来能够看到更多创新和突破性的成果。

#### 2.1.1 图像缺陷检测的分类与发展

从是否有监督信息（是否有异常样本或异常模式）的角度来看，视觉异常检测可以分为两个不同的研究方向：有监督和无监督的视觉异常检测。在本小节中，我们将主要回顾解决监督视觉异常检测问题的方法。一方面，有监督的视觉异常检测推动了所提出的方法的发展，以开发与人类视觉能力相似的计算机视觉模型。另一方面，对于大多数实际应用场景来说，异常样本或异常模式通常在形状、颜色和大小上都是可变的，并且它们没有稳定的统计规律，并在发展的过程中不断变化。所有这些都会使模型难以在变化的数据分布中捕获足够的统计信息或关于异常图像模式的显著特征。

根据不同的视觉检测精确颗粒度，视觉异常检测可以分为两类：图像级和像素级视

觉异常检测。其中，图像级检测通常只关注整个图像是否正常或异常，而像素级异常检测则需要进一步检测或定位图像中的异常区域。

此外，根据视觉异常检测研究的历史发展，（包括图像级和像素级）异常检测的文献大致可以分为两个阶段：深度学习之前和深度学习之后。在深度学习提出之前，视觉异常检测的研究重点在于开发异常检测策略或机制。主要的研究问题是：在通过手工获得图像的浅层特征（如灰度值、SIFT<sup>[1]</sup>和 HOG<sup>[2]</sup>）后，尝试基于统计或传统机器学习方法（如密度估计、单类分类和图像重建）开发不同的检测机制。它首先估计正常图像或图像特征的分布模型。然后，如果图像或其特征不符合相应的分布模型，它们将被识别为异常。

随着深度学习技术的发展，特别是在低级和高级计算机视觉任务中深度卷积神经网络取得巨大成功后<sup>[3]</sup>，相关研究人员逐渐将注意力转移到如何将深度卷积网络的强大表示能力与视觉异常检测问题相结合，并致力于开发端到端的检测方法。

其中，FCDD<sup>[10]</sup>是一种无监督方法，用于合成异常样本以训练一类分类（One-Class Classification, OCC）模型。这个方式也可以应用于其他一类分类方法。Venkataramanan 等人<sup>[11]</sup>提出了一种带有引导注意力的卷积对抗变分自编码器（CAVGA），该编码器可以同等地应用于有异常图像和无异常图像的情况。在无监督设置中，CAVGA 通过注意力扩展损失的引导使模型专注于图像的正常区域。在弱监督设置中，CAVGA 使用互补的引导注意力损失来最小化与图像异常区域相对应的注意力图，同时专注于正常区域。Božič 等人<sup>[12]</sup>研究了图像级监督信息、混合监督信息和像素级监督信息对同一深度学习框架内表面缺陷检测任务的影响。Božič 等人<sup>[12]</sup>发现，少量的像素级注释可以帮助模型实现与完全监督相当的性能。而 DevNet<sup>[13]</sup>则尝试使用少量的异常样本来实现细粒度的端到端可微分学习。Wan 等人<sup>[14]</sup>提出了一种用于训练具有不平衡数据分布的逻辑斯蒂诱导损失（LIS）和用于表征异常特征的异常捕获模块（ACM），以有效利用少量异常信息。

## 2.2 图像分割

图像分割是计算机视觉领域中最受欢迎的研究方向之一，并且是模式识别和图像理解的基础。图像分割技术的发展与许多学科和领域密切相关，例如自动驾驶<sup>[15]</sup>，智能医疗技术<sup>[16]</sup>，图像搜索引擎<sup>[18]</sup>，工业检测和增强现实。

图像分割将图像划分为具有不同特征的区域，并提取感兴趣的区域（ROI）。根据人类的视觉感知，这些区域是有意义的且不重叠的。图像分割有两个难点：（1）如何定义“有意义的区域”，由于视觉感知的不确定性和人类理解的多样性导致对象缺乏明确的定义，这使得图像分割成为一个不适定的问题；（2）如何有效地表示图像中的对象。数字图像由像素组成，这些像素可以根据它们的颜色，纹理和其他信息组合在一起来构成更大的集合。这些被称为“像素集”或“超像素”。这些低级特征反映了图像的局部属性，但很难通过这些局部属性获得全局信息（例如，形状和位置）。

自 20 世纪 70 年代以来, 图像分割一直受到计算机视觉研究人员的持续关注。经典的分割方法主要侧重于突出和获取单个图像中包含的信息, 这通常需要专业知识和人为干预。然而, 从图像中获得高级语义信息是很困难的。协同分割方法涉及从一组图像中识别出共同的对象, 这需要获取一定的先验知识。随着大规模精细标注图像数据集的丰富, 基于深度神经网络的图像分割方法逐渐成为研究热点。这些方法通常不需要详细的图像标注, 因此被归类为半监督或弱监督方法。

虽然图像分割研究已经取得了许多成果, 但仍存在许多挑战, 例如特征表示、模型设计和优化。特别是, 由于标注数据有限或稀疏、类别不平衡、过拟合、训练时间长和梯度消失等问题, 语义分割仍然面临诸多难题。

随着图像采集设备的不断发展, 图像细节的复杂性和物体间的差异 (例如尺度、姿态) 大大增加。低级特征 (例如颜色、亮度和纹理) 很难获得良好的分割结果, 而基于手动或启发式规则的特征提取方法无法满足当前图像分割的复杂需求, 这对图像分割模型的泛化能力提出了更高的要求。相应的, 深度学习算法被越来越多地应用于分割任务, 分割效果和性能得到了显著提升。原始的方法是将图像分成小块来训练神经网络, 然后对像素进行分类。由于神经网络的全连接层需要固定大小的图像, 因此采用了这种块分类算法<sup>[19]</sup>。

2015 年, Long 等人<sup>[20]</sup>提出了全卷积网络 (FCN), 用卷积代替全连接, 使得可以输入任意大小的图像, FCN 架构证明了神经网络可以进行端到端的语义分割训练, 为深度学习在语义分割中的应用奠定了基础。后续神经网络的发展大多基于 FCN 模型进行改进。下一节将介绍使用深度学习进行图像分割的主要技术和代表性模型。

### 2.2.1 编码器-解码器架构

在 FCN 之前, 卷积神经网络 (CNN) 在图像分类方面取得了良好效果, 例如 LeNet-5<sup>[21]</sup>、AlexNet<sup>[22]</sup>和 VGG<sup>[23]</sup>, 其输出层是图像的类别。然而, 语义分割需要在获得高级语义信息之后将高级特征映射回原始图像大小。对于这样的任务基于 FCN 的编码器-解码器架构十分有效。

在编码器阶段, 主要进行卷积和池化操作以提取包含语义信息的高维特征。卷积操作涉及将图像特定区域与不同的卷积核进行逐像素的乘法和求和, 然后通过激活函数变换获得特征图。池化操作涉及在特定区域 (池化窗口) 内进行采样, 然后使用某种采样统计量作为该区域的代表特征。在分割网络编码器中常用的骨干块是 VGG、Inception<sup>[24]</sup>和 ResNet<sup>[26]</sup>。

在解码器阶段, 通过高维特征向量生成语义分割掩码。将编码器提取的多级特征映射回原始图像的过程称为上采样。上采样的插值方法使用指定的插值策略在原始图像的像素之间插入新元素, 从而扩展图像的大小并实现上采样的效果。早期的上采样任务插值并不需要训练参数。而 FCN 采用反卷积进行上采样, 反卷积将原始卷积核的参数上下颠倒并水平翻转, 并在原始图像的元素之间及其周围填充空格。SegNet<sup>[27]</sup>则采

用反池化的上采样方法，反池化是 CNN 中最大池化的逆操作。在最大池化过程中，不仅要记录池化窗口的最大值，还要记录最大值的坐标位置；在反池化过程中，激活该位置的最大值，并将其他位置的值都设置为 0。Wang 等人<sup>[28]</sup>提出了一种密集上采样卷积（DUC），其核心思想是将特征图中的标签映射转换为具有多个通道的更小的标签映射。这种转换可以通过直接在输入特征图和输出标签图之间进行卷积来实现，而不需要在上采样过程中插值额外的值。

## 2.2.2 跳跃连接

跳跃连接或短路连接是为了改进粗糙的像素级定位而开发的。随着深度神经网络的训练，性能会随着深度的增加而降低，这是一个退化问题。为了缓解这个问题，ResNet 和 DenseNet<sup>[29]</sup>中提出了不同的跳跃连接结构。相比之下，U-Net<sup>[30]</sup>提出了一种新的长跳跃连接，如图 2-1 所示。

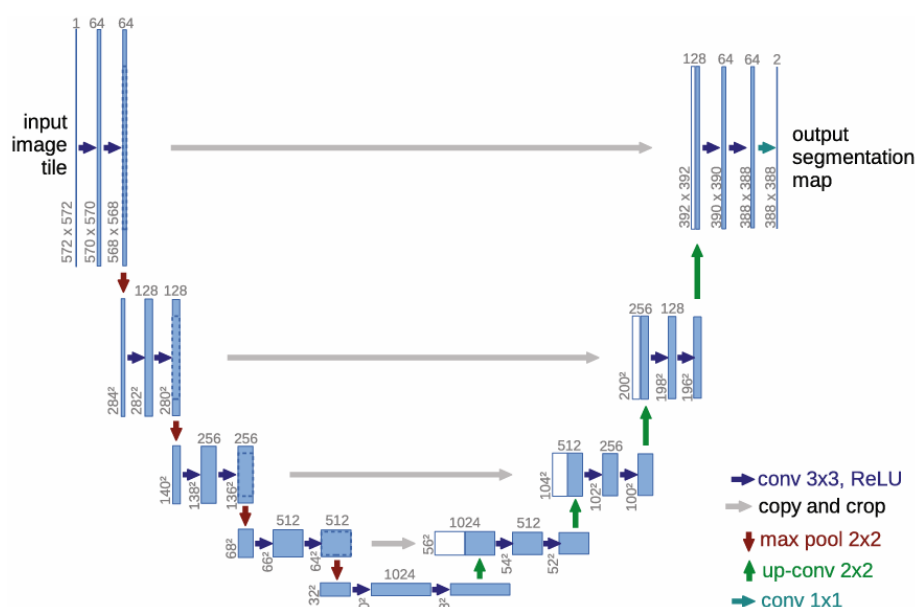


图 2-1 U-Net 结构图

U-Net 通过从编码器的层到解码器中相应层的跳跃连接和特征级联，获得图像的细粒度细节。它最初是为了解决生物显微图像分割问题而提出的，并自此在医学图像分割研究中得到广泛应用。U-Net 架构的设计灵感来自于全卷积网络（FCN），但它通过引入跳跃连接克服了 FCN 的一些限制。在 FCN 中，由于多次下采样操作，空间信息在网络的深层部分逐渐丢失，这可能导致分割结果中的细节信息不足。而 U-Net 通过跳跃连接将低层的空间信息与高层的语义信息结合起来，有效地解决了这个问题。

跳跃连接不仅有助于保持空间信息，还可以缓解梯度消失问题，使网络更容易训练。在反向传播过程中，梯度可以通过跳跃连接直接传递到较早的层，从而减少了梯度在传递过程中的消失。这使得 U-Net 等使用跳跃连接的网络能够在深度较大的情况下仍然保持良好的训练效果。



### 2.2.3 注意力机制

为了解决图像中不同区域之间的依赖关系，尤其是远距离区域，并获得它们的语义相关性，一些在自然语言处理（NLP）领域常用的方法已经被应用于计算机视觉，并在语义分割方面取得了良好的成果。注意力机制于 2014 年首次在计算机视觉领域提出。Google Mind 团队<sup>[31]</sup>采用循环神经网络（RNN）模型将注意力机制应用于图像分类，使注意力机制在图像处理任务中逐渐流行起来。

RNN 可以建立像素之间的短期依赖关系，连接像素，并按顺序处理它们，从而建立全局上下文关系。Visin 等人<sup>[32]</sup>基于 ReNet<sup>[33]</sup>提出了一种语义分割网络，每个 ReNet 层由四个 RNN 组成，这些 RNN 在图像的水平 and 垂直方向上扫描，以获得全局信息。

LSTM（长短期记忆）添加了一个新的功能来记录长期记忆，可以表示长距离依赖。Byeon 等人<sup>[34]</sup>使用 LSTM 实现了场景图像的像素级分割，证明了可以在二维 LSTM 模型中学习图像的纹理信息和空间模型参数。Liang 等人<sup>[35]</sup>基于图 LSTM 模型提出了一种语义分割模型，该模型将 LSTM 从序列数据或多维数据扩展到一般的图结构，进一步增强了全局上下文视觉特征。

自注意力机制则主要在编码器网络中使用，用于表示不同区域（像素）或特征图的不同通道之间的相关性。它计算单个样本所有位置之间键值对的加权和，以更新每个位置的特征。自注意力机制在图像分割方面取得了许多有影响力的成果，例如 PSANet<sup>[36]</sup>、DANet<sup>[37]</sup>、APCNet<sup>[38]</sup>、CARAFE<sup>[39]</sup>和 CARAFE++<sup>[40]</sup>。

2017 年，Vaswani 等人<sup>[41]</sup>提出了一种完全基于自注意力机制的深度学习神经网络 Transformer，完全摒弃了卷积和递归。此后，Transformer 及其变体（即 X-transformer）被应用于计算机视觉领域。利用 Transformer 的自注意力机制和 CNN 预训练模型，改进后的网络<sup>[42]</sup>也取得了一些突破。Dosovitskiy 等人<sup>[44]</sup>提出了一种视觉转换器（ViT），证明了 Transformer 可以替代 CNN 用于图像块序列的分类和预测。如图 2-2 所示，他们将图像划分为固定大小的块，排列图像块，并将块序列向量输入到由多头注意力层和多层感知器（MLP）组成的转换器编码器。

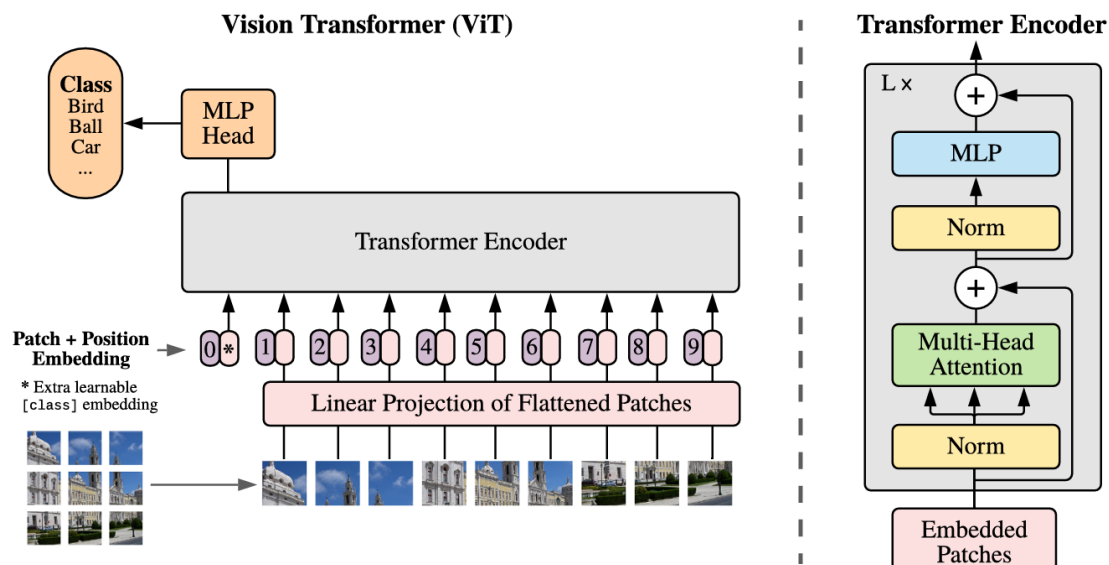


图 2-2 ViT 结构图

Liu 等人<sup>[45]</sup>开发了 Swin Transformer, 它在图像语义分割和实例分割方面取得了令人印象深刻的性能。Swin Transformer 采用了滑动窗口方法, 通过合并更深层次的图像块来构建不同层次的特征图, 在每个局部窗口中计算自注意力, 并在连续的 Swin Transformer 块中交替使用循环移位窗口来引入相邻非重叠窗口之间的跨窗口连接。如图 2-3 所示, Swin 转换器网络用移位窗口方法替换了转换器块中的标准多头自注意力 (MSA) 模块, 其他层保持不变。

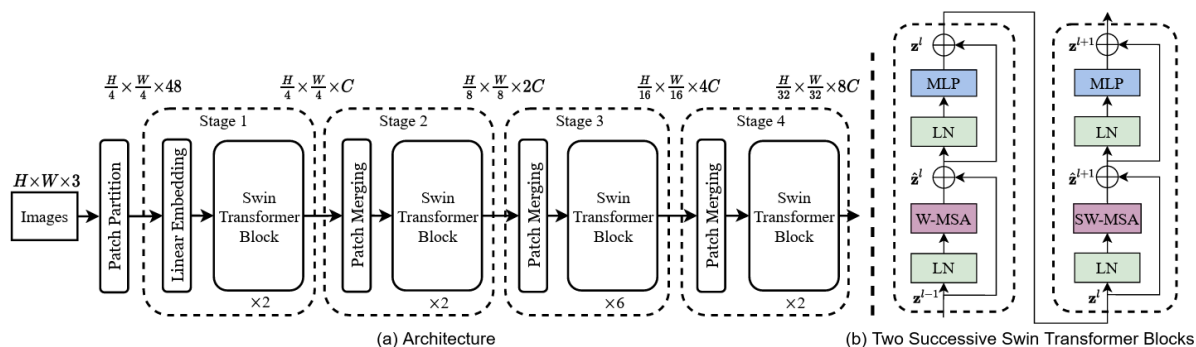


图 2-3 Swin Transformer 结构图

## 2.3 持续学习

深度学习中的一个重要问题是使神经网络能够增量地学习来自非平稳数据流的数据<sup>[46]</sup>。例如, 当深度神经网络在新任务或数据分布的样本上进行训练时, 它们往往会迅速失去先前获得的能力, 这种现象被称为灾难性遗忘<sup>[47]</sup>。与此形成鲜明对比的是, 人类和其他动物能够在不损害已经学到的技能的情况下增量地学习新技能<sup>[48]</sup>。持续学习, 也称为终身学习, 希望缩小生物智能和人工智能之间在增量学习能力方面的差距。近年来, 由于持续学习算法在医学诊断<sup>[49]</sup>、自动驾驶<sup>[50]</sup>或预测金融市场<sup>[51]</sup>等应用中的潜在实用性, 这一机器学习研究领域正在迅速扩大。

持续学习存在三种基本的有监督学习场景：

- (a) 在任务增量学习中，算法必须增量地学习一组可以清晰区分的任务；
- (b) 在领域增量学习中，算法必须在不同的上下文中学习相同类型的问题；
- (c) 在类别增量学习中，算法必须增量地学习区分越来越多的对象或类别。

在下一小节，本文将更详细地介绍这三种场景，并指出了与每种场景相关的不同挑战。同时也会回顾使用持续学习来优化深度神经网络的现有策略，并简单比较这些不同的策略对于每个场景的适用性。

### 2.3.1 持续学习的三个场景

在经典的机器学习中场景，算法可以同时访问所有训练数据。而在持续学习中，数据则是按顺序或分步骤到达的，并且数据的基础分布在随时间变化。

第一个持续学习场景是“任务增量学习”(Task-incremental Learning, 简称 Task-IL)。任务增量学习的定义是，在测试时始终清楚需要执行哪个任务。在实践中，这可能意味着任务身份是明确提供的，或者任务之间是可以明确区分的。在这种场景中，可以使用具有特定任务模组的模型进行训练（例如，每个任务都有一个单独的输出层），甚至可以为每个要学习的任务分配一个完全独立的网络，在后一种情况下，根本不存在灾难性遗忘的问题。因此，任务递增学习的挑战不在于简单地防止灾难性遗忘，而是找到有效的方式来共享跨任务的知识，以优化性能和计算复杂度之间的关系。并希望能够进一步地利用从不同任务中学习到的信息来提高其他任务的性能<sup>[52]</sup>。任务递增学习在现实中的例子是学习不同的运动或演奏不同的乐器，通常在学习过程中学习者总是很清楚应该学习哪种运动或演奏哪种乐器。

第二种场景称为“领域增量学习”（或 Domain-Incremental Learning, 简称 Domain-IL）。在这个场景中，问题的结构总是相同的，但上下文或输入分布会发生变化。与任务增量学习类似，这个场景也可以描述为算法必须增量地学习一组不同领域的相同任务，但关键的区别在于算法不知道样本属于哪个领域。然而，模型并不一定需要识别当前任务所在的领域，因为每个领域上的任务都会使用相同的类别输出。然而，在这个场景中，算法只有在识别出任务所在的领域之后，才有可能使用特定任务的组件<sup>[53]</sup>，这相较于 Task-IL 所在的场景而言明显不那么合适。因此，在领域增量学习中，不可能通过为每一个任务设计不同的模组来防止遗忘。因此，在 Domain-IL 中，缓解灾难性遗忘仍然是一个重要的未解决问题。这个场景在现实中的例子包括在可变光照条件下逐步学习识别物体<sup>[56]</sup>或在不同天气条件下学习自动驾驶<sup>[57]</sup>。

第三种持续学习场景是“类别增量学习”（或 Class-Incremental Learning, 简称 Class-IL）。在这个场景中算法必须增量地学习区分越来越多的对象或类别。这个场景常用的设置是遇到一系列基于分类的任务，其中每个任务包含不同的类别，算法必须学习区分所有类别<sup>[58]</sup>。在这种情况下，任务识别对于解决问题是必要的，因为它决定了当前

样本可能属于哪些可能的类别。换句话说，算法应该既能解决每个单独的任务，又能识别样本属于哪个任务。例如，模型可能首先学习区分猫和狗，然后学习区分牛和马；而在任务增量学习中，模型不需要区分不同领域中遇到的动物（例如猫和牛），但在类别递增学习中这是必需的。这个场景中的一个重要挑战是学习区分没有一起观察到的类别，尤其是当我们只有有限的资源来存储先前任务的信息时，这对深度神经网络来说是十分难以解决的问题。

### 2.3.2 持续学习的三种方案

我们将现有的持续学习工作大致分为三类。在第一类中，我们讨论基于重要性加权参数正则化的方法，该方法使神经网络的学习过程减少遗忘。在第二类中，我们讨论基于数据重放的方法，这些方法要么存储以前任务中的示例，要么学习从先前观察到的数据分布生成新示例。在第三类中，我们讨论随时间增长改变神经网络架构的方法。

基于正则化的方法。这类方法的共同关键思想是识别在学习过去经验中起重要作用的参数。然后，这些参数在未来更新中受到保护，而不重要的参数则进一步训练以学习新任务。该类方法通常通过向目标添加一个正则化项来解决灾难性遗忘问题，该正则化项会惩罚神经网络输入输出函数的变化。在 Li 和 Hoiem 的研究中<sup>[59]</sup>，他们使用一种知识蒸馏的形式。而 Hinton<sup>[60]</sup>，鼓励先前任务网络和当前网络在新任务数据上的预测结果相似。类似地，Jung 等人<sup>[61]</sup>对最终隐藏激活之间的 L2 距离进行了正则化，而不是使用知识蒸馏惩罚。这两种正则化方法都旨在通过使用旧任务的参数存储或计算额外的激活函数来保留旧任务的输入输出的一部分映射。这使得功能性方法在计算上变得昂贵，因为它需要对每个新数据点通过旧任务的网络进行一次前向传播。

基于数据重放的方法。这类方法假设我们只有有限的内存空间来存储以前任务的示例，而模型将使用这些有限示例来防止在学习新任务时出现灾难性遗忘。在这样的场景下，生成模型显得十分有益，因为生成模型能够从学习的数据分布中抽取伪示例，为基于记忆的模型提供了一种替代的采样方法。例如 Robins<sup>[62]</sup>发现将新的信息与先前内部生成模式交织在一起，有助于巩固现有知识，而无需显式存储训练样本。Draelos 等人<sup>[63]</sup>也使用伪排练方法对自编码器进行增量训练，在数据重放过程中利用编码器的输出统计信息生成解码器的输入。然而，与上述大多数方法类似，伪排练方法的使用仅在两个相对低复杂度的数据集上进行了严格评估，例如 MNIST 和 Street View House Number。因此，这种生成方法是否能扩展到更复杂的领域，这是一个值得探讨的问题。

基于动态神经网络的方法。到目前为止所讨论的方法都假设有一个无限容量的网络来学习新任务。然而在实际场景中，神经网络的有限容量会限制其随时间学习新任务的能力。关于动态神经网络的各种工作<sup>[64]</sup>旨在解决这个问题。这些方法从一个简化的架构开始，并在需要时逐步增加新组件以扩展网络，从而在当前任务上获得满意的性能。在<sup>[66]</sup>中，神经网络模型通过增加一个称为列的单独子网络来进行扩展，该子网络负责学习

新任务。先前学习任务中开发的列表示保持固定，以避免灾难性干扰。然后，这些表示作为新添加的列的额外输入进行馈送。然而，这种方法存在扩展性问题，因为每次需要学习新任务时都会添加一个新列。由于需要对先前添加的列进行计算以将信号传递到新列，因此此类网络中的推理变得越来越困难。最近，Jaehong<sup>[66]</sup>提出了一种创新的方法，该方法随时间仅亚线性地增长网络的规模。

1.1 二级标题（黑体四号，左对齐，段前距为 12 磅）

内容<sup>[1]</sup>（内容首行缩进两格，宋体小四号；引文在引用处的右上角标注，五号字）

1.1.1 三级标题（黑体小四号，左对齐，段前距为 12 磅）

内容<sup>[2-15]</sup>（内容首行缩进两格，宋体小四号；引文在引用处的右上角标注，五号字）

表 1.1 表名（编号与表名相距半角状态下的一个空格，表正上方，宋体五号粗体）

---

内容（宋体五号）
----------

---

---

备注：表格采用三线表，不允许断页。（宋体五号）

1.2 二级标题（黑体四号，左对齐，段前距为 12 磅）

1.2.1 三级标题（黑体小四号，左对齐，段前距为 12 磅）

内容（首行缩进两格，宋体小四号）

2 一级标题（换页并首行居中，黑体小二号，段前距、段后距分别为 12 磅）

2.1 二级标题（黑体四号，左对齐，段前距为 12 磅）

2.1.1 三级标题（黑体小四号，左对齐，段前距为 12 磅）

内容（首行缩进两格，宋体小四号）

$A = 0.235B + 2.68C$ （另起行缩进两格，公式编辑器编辑，五号字）（2-1）（五号字，靠右编排）

2.2 二级标题（黑体四号，左对齐，段前距为 12 磅）

2.2.1 三级标题（黑体小四号，左对齐，段前距为 12 磅）

内容<sup>[2,8]</sup>（内容首行缩进两格，宋体小四号；引文在引用处的右上角标注，五号字）

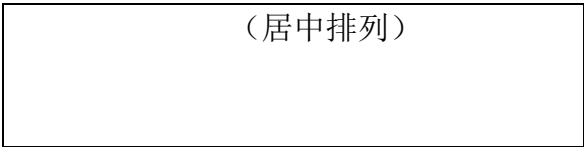


图 2-1 图名（编号与图名相距半角状态下的一个空格，图正下方，宋体五号粗体）

### 3 一级标题（换页并首行居中，黑体小二号，段前距、段后距分别为 12 磅）



## 参考文献

- [1] Lowe D G. Object recognition from local scale-invariant features[C]//Proceedings of the seventh IEEE international conference on computer vision. Ieee, 1999, 2: 1150-1157.
- [2] Dalal N, Triggs B. Histograms of oriented gradients for human detection[C]//2005 IEEE computer society conference on computer vision and pattern recognition (CVPR'05). Ieee, 2005, 1: 886-893.
- [3] Zhang Y, Tian Y, Kong Y, et al. Residual dense network for image restoration[J]. IEEE transactions on pattern analysis and machine intelligence, 2020, 43(7): 2480-2495.
- [4] Yang J, Qi Z, Shi Y. Learning to incorporate structure knowledge for image inpainting[C]//Proceedings of the AAAI conference on artificial intelligence. 2020, 34(07): 12605-12612.
- [5] Xie S, Tu Z. Holistically-nested edge detection[C]//Proceedings of the IEEE international conference on computer vision. 2015: 1395-1403.
- [6] Li K, Tian Y, Wang B, et al. Bi-directional pyramid network for edge detection[J]. Electronics, 2021, 10(3): 329.
- [7] Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition[J]. arXiv preprint arXiv:1409.1556, 2014.
- [8] 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.
- [9] Sun R, Zhu X, Wu C, et al. Not all areas are equal: Transfer learning for semantic segmentation via hierarchical region selection[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2019: 4360-4369.
- [10] Liznerski P, Ruff L, Vandermeulen R A, et al. Explainable deep one-class classification[J]. arXiv preprint arXiv:2007.01760, 2020.
- [11] Venkataramanan S, Peng K C, Singh R V, et al. Attention guided anomaly localization in images[C]//European Conference on Computer Vision. Cham: Springer International Publishing, 2020: 485-503.
- [12] Božič J, Tabernik D, Skočaj D. Mixed supervision for surface-defect detection: From weakly to fully supervised learning[J]. Computers in Industry, 2021, 129: 103459.

- [13]Pang G, Ding C, Shen C, et al. Explainable deep few-shot anomaly detection with deviation networks[J]. arXiv preprint arXiv:2108.00462, 2021.
- [14]Wan Q, Gao L, Li X. Logit inducing with abnormality capturing for semi-supervised image anomaly detection[J]. IEEE Transactions on Instrumentation and Measurement, 2022, 71: 1-12.
- [15]Kabiraj A, Pal D, Ganguly D, et al. Number plate recognition from enhanced super-resolution using generative adversarial network[J]. Multimedia Tools and Applications, 2023, 82(9): 13837-13853.
- [16]Jin B, Cruz L, Gonçalves N. Deep facial diagnosis: deep transfer learning from face recognition to facial diagnosis[J]. IEEE Access, 2020, 8: 123649-123661.
- [17]Zhao M, Liu Q, Jha A, et al. VoxelEmbed: 3D instance segmentation and tracking with voxel embedding based deep learning[C]//Machine Learning in Medical Imaging: 12th International Workshop, MLMI 2021, Held in Conjunction with MICCAI 2021, Strasbourg, France, September 27, 2021, Proceedings 12. Springer International Publishing, 2021: 437-446.
- [18]Yao T, Qu C, Liu Q, et al. Compound figure separation of biomedical images with side loss[C]//Deep Generative Models, and Data Augmentation, Labelling, and Imperfections: First Workshop, DGM4MICCAI 2021, and First Workshop, DALI 2021, Held in Conjunction with MICCAI 2021, Strasbourg, France, October 1, 2021, Proceedings 1. Springer International Publishing, 2021: 173-183.
- [19]Li H, Zhao R, Wang X. Highly efficient forward and backward propagation of convolutional neural networks for pixelwise classification[J]. arXiv preprint arXiv:1412.4526, 2014.
- [20]Shelhamer E, Long J, Darrell T. Fully convolutional networks for semantic segmentation[J]. IEEE Trans. Pattern Anal. Mach. Intell., 2017, 39(4): 640-651.
- [21]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.
- [22]Krizhevsky A, Sutskever I, Hinton G E. ImageNet classification with deep convolutional neural networks[J]. Communications of the ACM, 2017, 60(6): 84-90.
- [23]Karen S, Zisserman A. Very deep convolutional networks for large-scale image recognition. arXiv 2014[J]. arXiv preprint arXiv:1409.1556, 2014.
- [24]Szegedy C, Liu W, Jia Y, et al. Going deeper with convolutions[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2015: 1-9.
- [25]Szegedy C, Vanhoucke V, Ioffe S, et al. Rethinking the inception architecture for computer vision[C]//Proceedings of the IEEE conference on computer vision and

- pattern recognition. 2016: 2818-2826.
- [26]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.
- [27]Badrinarayanan V, Kendall A, Cipolla R. Segnet: A deep convolutional encoder-decoder architecture for image segmentation[J]. IEEE transactions on pattern analysis and machine intelligence, 2017, 39(12): 2481-2495.
- [28]Wang P, Chen P, Yuan Y, et al. Understanding convolution for semantic segmentation[C]//2018 IEEE winter conference on applications of computer vision (WACV). Ieee, 2018: 1451-1460.
- [29]Maaten L, Huang G, Liu Z, et al. Densely connected convolutional networks[C]. CVPR, 2017.
- [30]Ronneberger O, Fischer P, Brox T. U-net: Convolutional networks for biomedical image segmentation[C]//Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015: 18th International Conference, Munich, Germany, October 5-9, 2015, Proceedings, Part III 18. Springer International Publishing, 2015: 234-241.
- [31]Mnih V, Heess N, Graves A. Recurrent models of visual attention[J]. Advances in neural information processing systems, 2014, 27.
- [32]Visin F, Ciccone M, Romero A, et al. Reseg: A recurrent neural network-based model for semantic segmentation[C]//Proceedings of the IEEE conference on computer vision and pattern recognition workshops. 2016: 41-48.
- [33]Visin F, Kastner K, Cho K, et al. Renet: A recurrent neural network based alternative to convolutional networks[J]. arXiv preprint arXiv:1505.00393, 2015.
- [34]Byeon W, Breuel T M, Raue F, et al. Scene labeling with lstm recurrent neural networks[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2015: 3547-3555.
- [35]Liang X, Shen X, Feng J, et al. Semantic object parsing with graph lstm[C]//Computer Vision–ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11–14, 2016, Proceedings, Part I 14. Springer International Publishing, 2016: 125-143.
- [36]Zhao H, Zhang Y, Liu S, et al. Psanet: Point-wise spatial attention network for scene parsing[C]//Proceedings of the European conference on computer vision (ECCV). 2018: 267-283.
- [37]Fu J, Liu J, Tian H, et al. Dual attention network for scene

- segmentation[C]//Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. 2019: 3146-3154.
- [38]He J, Deng Z, Zhou L, et al. Adaptive pyramid context network for semantic segmentation[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2019: 7519-7528.
- [39]Wang J, Chen K, Xu R, et al. Carafe: Content-aware reassembly of features[C]//Proceedings of the IEEE/CVF international conference on computer vision. 2019: 3007-3016.
- [40]Wang J, Chen K, Xu R, et al. Carafe++: Unified content-aware reassembly of features[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2021, 44(9): 4674-4687.
- [41]Vaswani A, Shazeer N, Parmar N, et al. Attention is all you need[J]. Advances in neural information processing systems, 2017, 30.
- [42]Weissenborn D, Täckström O, Uszkoreit J. Scaling autoregressive video models[J]. arXiv preprint arXiv:1906.02634, 2019.
- [43]Cordonnier J B, Loukas A, Jaggi M. On the relationship between self-attention and convolutional layers[J]. arXiv preprint arXiv:1911.03584, 2019.
- [44]Dosovitskiy A, Beyer L, Kolesnikov A, et al. An image is worth 16x16 words: Transformers for image recognition at scale[J]. arXiv preprint arXiv:2010.11929, 2020.
- [45]Liu Z, Lin Y, Cao Y, et al. Swin transformer: Hierarchical vision transformer using shifted windows[C]//Proceedings of the IEEE/CVF international conference on computer vision. 2021: 10012-10022.
- [46]Chen Z, Liu B. Lifelong machine learning[M]. San Rafael: Morgan & Claypool Publishers, 2018.
- [47]French R M. Catastrophic forgetting in connectionist networks[J]. Trends in cognitive sciences, 1999, 3(4): 128-135.
- [48]Kudithipudi D, Aguilar-Simon M, Babb J, et al. Biological underpinnings for lifelong learning machines[J]. Nature Machine Intelligence, 2022, 4(3): 196-210.
- [49]Lee C S, Lee A Y. Clinical applications of continual learning machine learning[J]. The Lancet Digital Health, 2020, 2(6): e279-e281.
- [50]Shaheen K, Hanif M A, Hasan O, et al. Continual learning for real-world autonomous systems: Algorithms, challenges and frameworks[J]. Journal of Intelligent & Robotic Systems, 2022, 105(1): 9.
- [51]Philps D, Weyde T, Garcez A A, et al. Continual learning augmented investment

- decisions[J]. arXiv preprint arXiv:1812.02340, 2018.
- [52] Lopez-Paz D, Ranzato M A. Gradient episodic memory for continual learning[J]. Advances in neural information processing systems, 2017, 30.
- [53] Aljundi R, Chakravarty P, Tuytelaars T. Expert gate: Lifelong learning with a network of experts[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2017: 3366-3375.
- [54] Von Oswald J, Henning C, Grewe B F, et al. Continual learning with hypernetworks[J]. arXiv preprint arXiv:1906.00695, 2019.
- [55] Wortsman M, Ramanujan V, Liu R, et al. Supermasks in superposition[J]. Advances in Neural Information Processing Systems, 2020, 33: 15173-15184.
- [56] Lomonaco V, Maltoni D. Core50: a new dataset and benchmark for continuous object recognition[C]//Conference on robot learning. PMLR, 2017: 17-26.
- [57] Mirza M J, Masana M, Possegger H, et al. An efficient domain-incremental learning approach to drive in all weather conditions[C]//Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. 2022: 3001-3011.
- [58] Shin H, Lee J K, Kim J, et al. Continual learning with deep generative replay[J]. Advances in neural information processing systems, 2017, 30.
- [59] Li Z, Hoiem D. Learning without forgetting[J]. IEEE transactions on pattern analysis and machine intelligence, 2017, 40(12): 2935-2947.
- [60] Hinton G, Vinyals O, Dean J. Distilling the knowledge in a neural network[J]. arXiv preprint arXiv:1503.02531, 2015.
- [61] Jung H, Ju J, Jung M, et al. Less-forgetting learning in deep neural networks[J]. arXiv preprint arXiv:1607.00122, 2016.
- [62] Robins A. Catastrophic forgetting, rehearsal and pseudorehearsal[J]. Connection Science, 1995, 7(2): 123-146.
- [63] Draelos T J, Miner N E, Lamb C C, et al. Neurogenesis deep learning: Extending deep networks to accommodate new classes[C]//2017 international joint conference on neural networks (IJCNN). IEEE, 2017: 526-533.
- [64] Cortes C, Gonzalvo X, Kuznetsov V, et al. Adanet: Adaptive structural learning of artificial neural networks[C]//International conference on machine learning. PMLR, 2017: 874-883.
- [65] Yoon J, Yang E, Lee J, et al. Lifelong learning with dynamically expandable networks[J]. arXiv preprint arXiv:1708.01547, 2017.
- [66] Rusu A A, Rabinowitz N C, Desjardins G, et al. Progressive neural networks[J]. arXiv preprint arXiv:1606.04671, 2016.

- 
- [67] Carta A, Pellegrini L, Cossu A, et al. Avalanche: A pytorch library for deep continual learning[J]. Journal of Machine Learning Research, 2023, 24(363): 1-6.
- [68] Kirkpatrick J, Pascanu R, Rabinowitz N, et al. Overcoming catastrophic forgetting in neural networks[J]. Proceedings of the national academy of sciences, 2017, 114(13): 3521-3526.
- [69] Zenke F, Poole B, Ganguli S. Continual learning through synaptic intelligence[C]//International conference on machine learning. PMLR, 2017: 3987-3995

致谢（换页，黑体小二号居中，段前距、段后距分别为 12 磅）

内容（首行缩进两格，小四号字）

XXX  
XXXX 年 XX 月 XX 日