2．**项目的研究内容、研究目标，以及拟解决的关键科学问题**（此部分为重点阐述内容）**；**

**2.1 研究内容**

本项目围绕集成电路封装工艺智能化质量检测的工程迫切性需求，开展以集成电路关键元器件为研究对象，在低照度低分辨率数字相机处理流水线环境下，针对数据不全面和正负样本数据不均衡等普遍现象，从提升数据质量和数量角度出发，研究基于领域自适应和零样本学习的数据扩充技术；从模型轻量化和轻量化学习策略角度出发，优化基于在线学习的集成电路关键元器件智能检测模型，从而完成元器件封装状态及封装质量、贴片标识等自适应高速识别任务，进一步将智慧生产思维赋能于实际工业生产环节，助力集成电路面向无人化“灯塔工厂”发展及普及。本项目具体研究内容包括：

**（1）低照度低分辨率数字相机处理流水线下样本不全面不均衡数据鸿沟挑战**

在实际工业场景的芯片封装产品线中，由于封闭车间内自然光照匮乏，人工光照覆盖范围小、照射区域不均衡、光照强度弱，拍摄采集到的芯片封装图像样本具有低照度低分辨率的特点，样本质量对机器视觉模型的检测识别带来了较大的困难。此外，产品往往具有良率高、封装方式多样、流程繁复等特点，由此导致缺陷出现频率较低、缺陷类型多样、不同缺陷的出现频率和对产品质量影响程度差别很大，这也极大地影响了机器视觉缺陷检测模型的训练和应用。一方面，有缺陷的样本相对于无缺陷样本，数量稀缺、获取困难，无法获取足够的训练样本难以支持深度学习模型的训练。同时，芯片类型多样、封装方式多样、缺陷问题多样，现有的缺陷数据样本不能够覆盖潜在的可能缺陷类型，从而使得样本多样性较差，制约了模型的泛化能力。另一方面，在现有的缺陷类型中，不同类型缺陷出现的频率不同，有些低频关键缺陷对芯片封装质量影响较大，但其在训练和测试样本中出现次数少，表现为了不同类别间的样本不均衡问题，这会导致模型忽略低频关键缺陷，从而影响缺陷检测精度。上述因素的并存所导致的芯片封装流水线下缺陷图像质量和数量问题对机器视觉缺陷检测方法带来了数据鸿沟挑战。本项目将围绕上述问题，开展低照度图像感知质量提升方法研究、生成对抗式多样化缺陷图像生成模型研究和缺陷图像样本间均衡性自适应模型，解决数据鸿沟问题。主要研究内容有：

**① 低照度图像感知质量提升方法研究**：从模拟初级视觉皮层图像感知机理角度出发，研究基于Retinex图像增强物理模型；借助具备模仿人类推理过程的零次学习理论，制定零参考损失、深度预设和优化策略；应用深度学习于Retinex仿生模型，设计基于双边网络的图像增强模型，并通过设计约束条件和先验的损失函数，纠正不均衡光照、提高泛化能力和去除未知噪声干扰，从而提升集成电路关键元器件低照度图像感知质量。

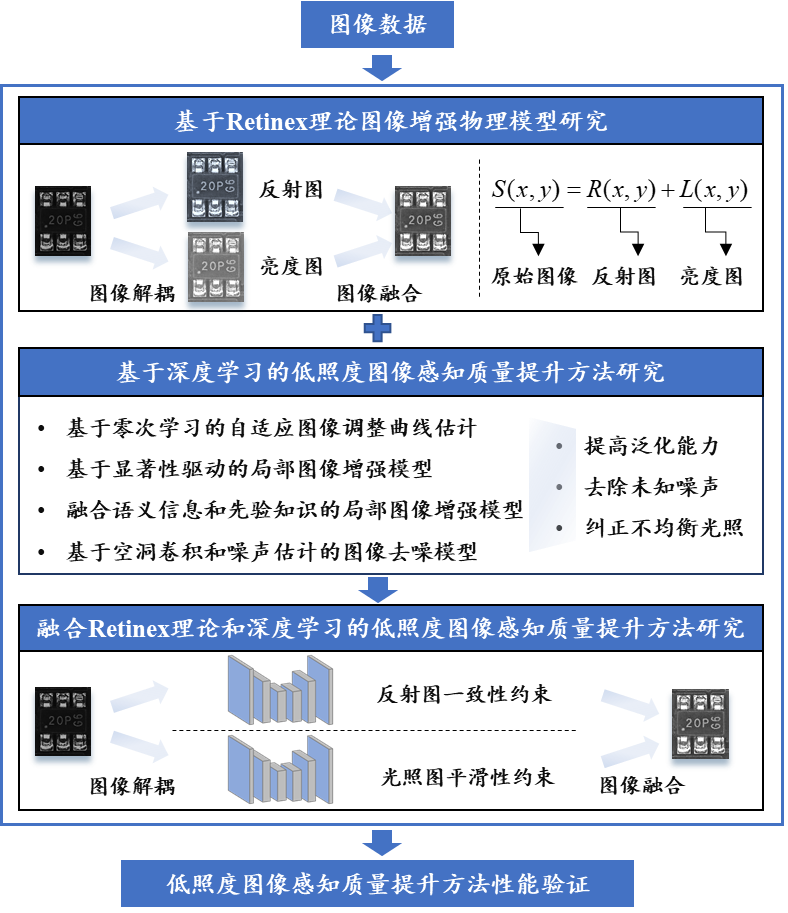


图1 低照度图像感知质量提升方法

**② 生成对抗式多样化缺陷图像生成模型研究**：针对小样本挑战，在本项目的研究中，拟构建基于生成对抗网络的小样本零样本缺陷图像生成模型；采用条件生成对抗网络技术，基于现有缺陷图像生成相似的缺陷图像从而缓解数据稀缺问题；采用生成对抗网络与变分自编码器相结合的技术，利用无缺图像样本与随机生成噪声块构建人工合成缺陷图像，通过对现有缺陷进行随机组合和引入噪声来生成多样化缺陷图像，提升样本多样性。

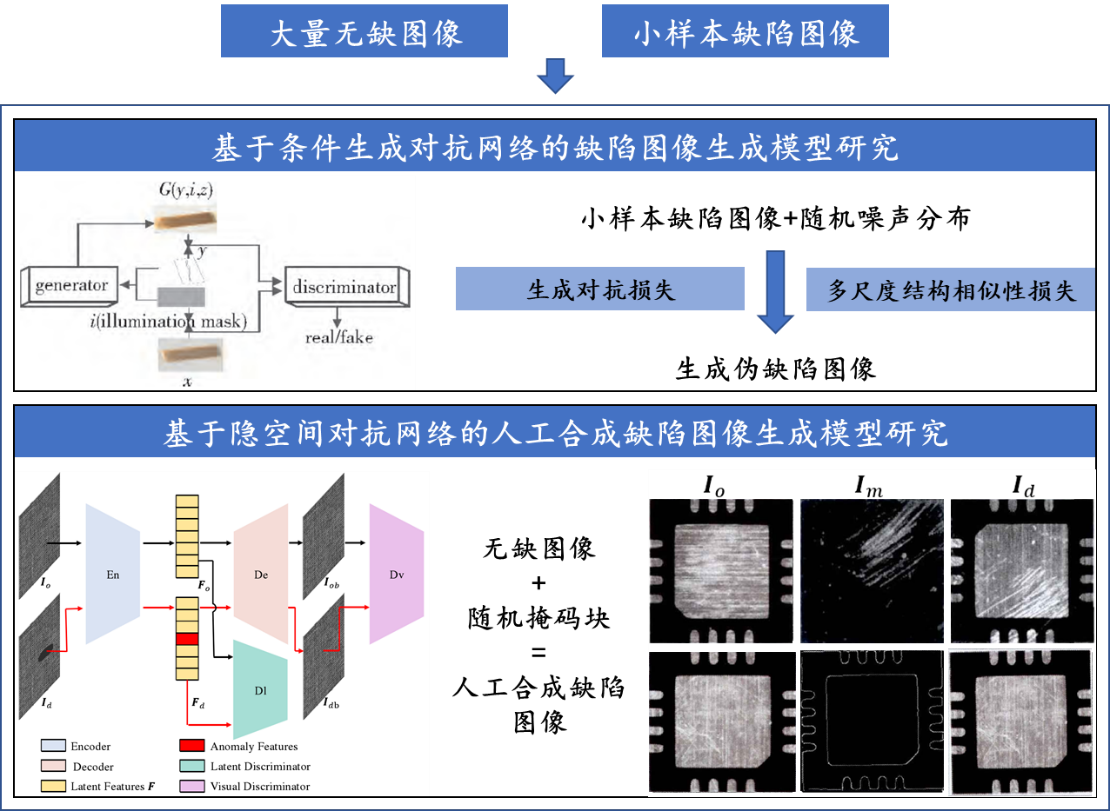
****

图2 生成对抗式多样化缺陷图像生成模型研究

**③ 缺陷图像样本间均衡性自适应模型研究**：针对样本不均衡问题，本项目拟构建基于样本重采样和注意力机制的缺陷图像均衡性自适应模型；引入样本重采样技术，通过对低频缺陷图像的多次采样来实现样本间的均衡，同时拟引入自适应注意力机制，通过对样本加权和施加多粒度损失约束来引导模型关注低频关键缺陷模式，从而克服样本不均衡问题。

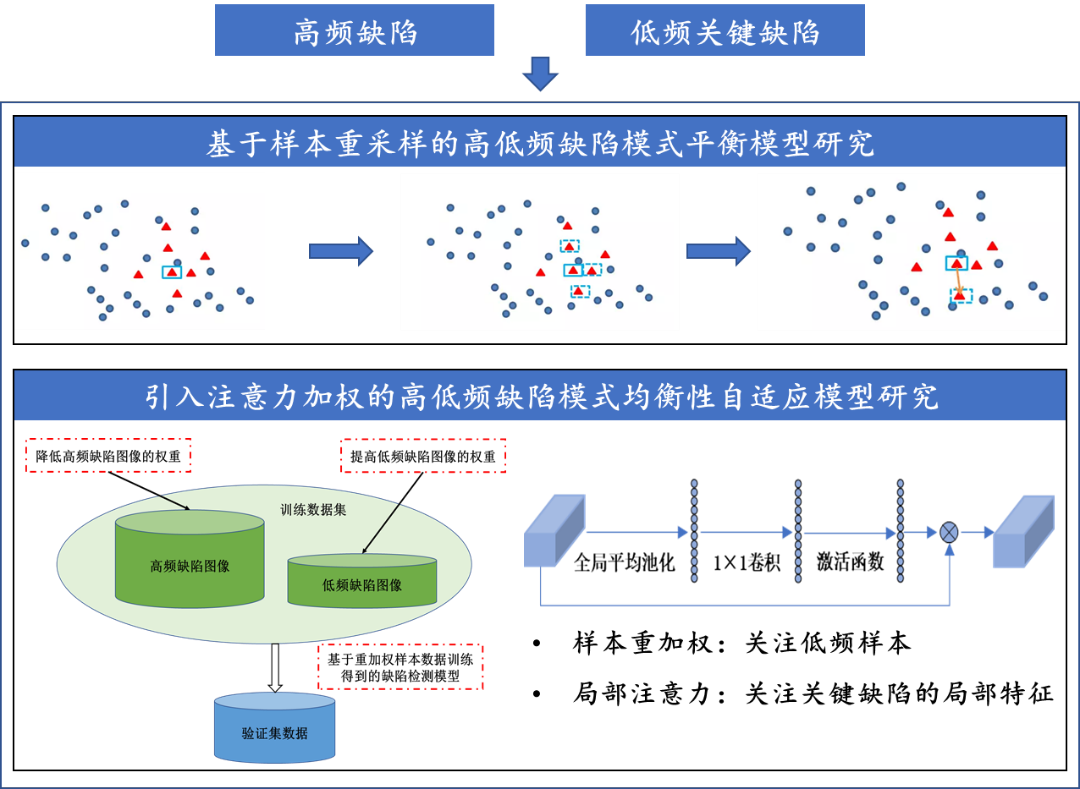
****

图3 缺陷图像样本间均衡性自适应模型研究

**（2）基于集成电路封装缺陷检测的大模型轻量化研究**

目前深度神经网络在各个工业领域中取得了优异的成果，其根本上还是在于有着海量的数据、庞大的模型、高性能的计算的加持。而这也在某种程度上限制了深度学习的进一步发展和普及。半导体封装缺陷检测的小样本以及样本不均衡特性导致其检测模型只有提高参数量和计算量才能较好地提取到缺陷样本特征从而区分有缺陷的样本和正常的样本。庞大的模型和高性能的计算需要拥有大量存储空间和优异计算性能的硬件平台，然而半导体封装生产线上往往缺乏这样的条件，这也导致将庞大的检测模型应用于半导体封装生产线场景中成为了巨大的挑战。因此落地于半导体封装生产线上的缺陷检测模型必须在保证检测精度的前提下尽量降低参数量、计算量，减少网络推理时间。本项目将围绕模型轻量化问题从模型本身和检测策略方面出发进行研究。主要研究内容如下：

**① 硬件加速反馈的精细化通道剪枝的超轻量模型研究**：有机结合硬件加速反馈的模型轻量化是大型神经网络部署在半导体封装产业线上的前提。研究引入硬件性能预测器将硬件加速信息反馈到剪枝流程中，从而对给定的硬件平台和资源约束，提高模型轻量化效率。同时在剪枝算法与模型结构层面上对剪枝进行优化，研究自动可学的剪枝优化算法，在低秩分解后的通道空间进行细粒度的高效剪枝。

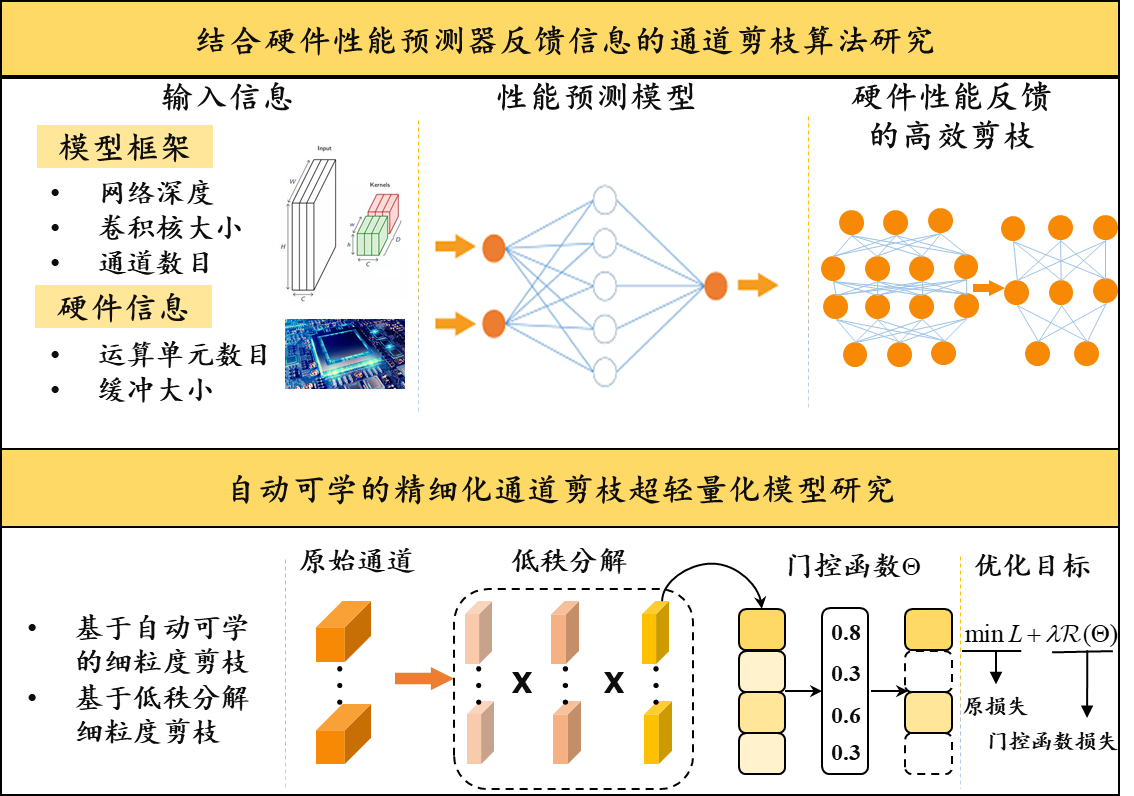
****

图4 硬件加速反馈的精细化通道剪枝的超轻量模型研究

**② 跨步推理与多点退出的快速推理检测策略研究**：动态跳过某些不重要的计算和及时退出网络停止计算是提高半导体封装缺陷检测推理速度达到实时检测要求的有效手段。研究跨步推理和多点退出策略，引入最优辅助网络帮助判别模型计算过程中哪些步骤不重要可以直接跳过，检测任务运算到哪一步已经满足精度要求可以退出，从而加速推理速度。

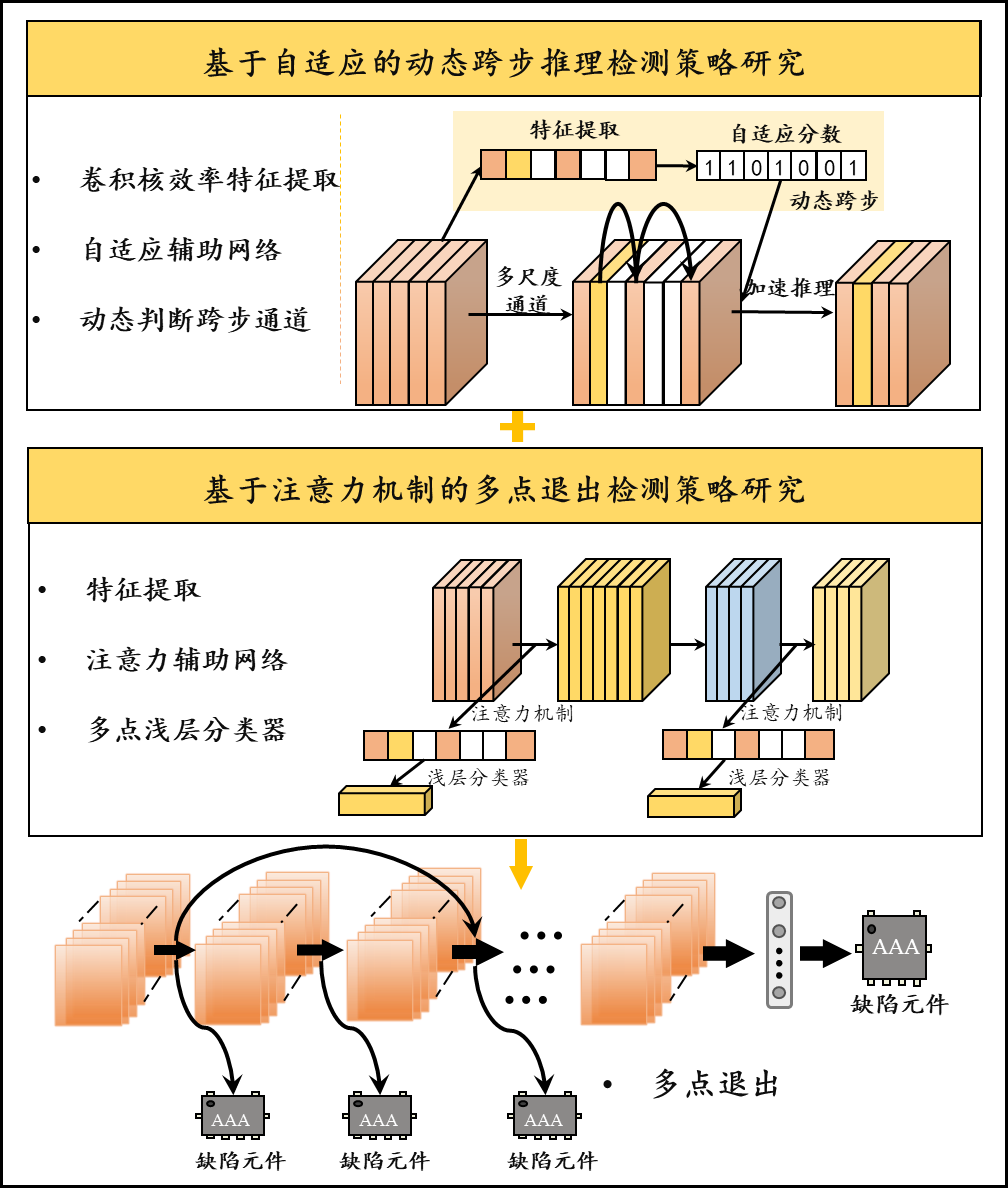


图5 跨步推理与多点退出的快速推理检测策略研究

**（3）基于感知数据驱动的集成电路关键元器件在线检测智能分析研究**

集成电路中关键元器件常见的封装工艺包括插式封装和熔焊封装，其封装状态和封装质量直接或间接影响着整个终端产品的质量水平。本研究以集成电路关键元器件贴片为研究对象，重点解决集成电路封装工艺中面临的“卡脖子”问题，如：贴片反贴、极反识别，贴片多件、缺件/缺锡识别，贴片偏移量化等封装状态和封装质量识别与检测难点。另一方面，贴片表面的字母、数字和图像符号组成的标识携带了一些重要的生产信息，如：厂商号、版本号、批次号及识别代码等，这些信息被用于电子工业的产品追踪、质量控制、产品召回管理、生产信息管理，是终端产品生命周期中跟踪追溯的重要凭据。然而由于贴片尺寸较小、标识印刷模糊、外观多样性高、受到焊接材料干扰等产品自身特性的难点，研究基于自适应高速辨识标识问题成为极其具有应用价值的科学问题。本项目将围绕上述实际工程应用问题，开展集成电路贴片封装状态与封装质量检测方法和贴片标识自适应高速辨识方法研究，从而应用于流水线作业方式下集成电路贴片在线智能检测任务。主要研究内容有：

**① 集成电路贴片封装状态与封装质量检测方法研究**：研究基于跨网络层信息交互和多尺度信息融合的深度特征提取模型，从而提取图像数据中感兴趣区域；引入级联注意力机制，完成贴片中关键位置准确定位及分类，并通过设计封装状态和封装质量的评价指标，进而实现贴片封装状态与封装质量检测任务。

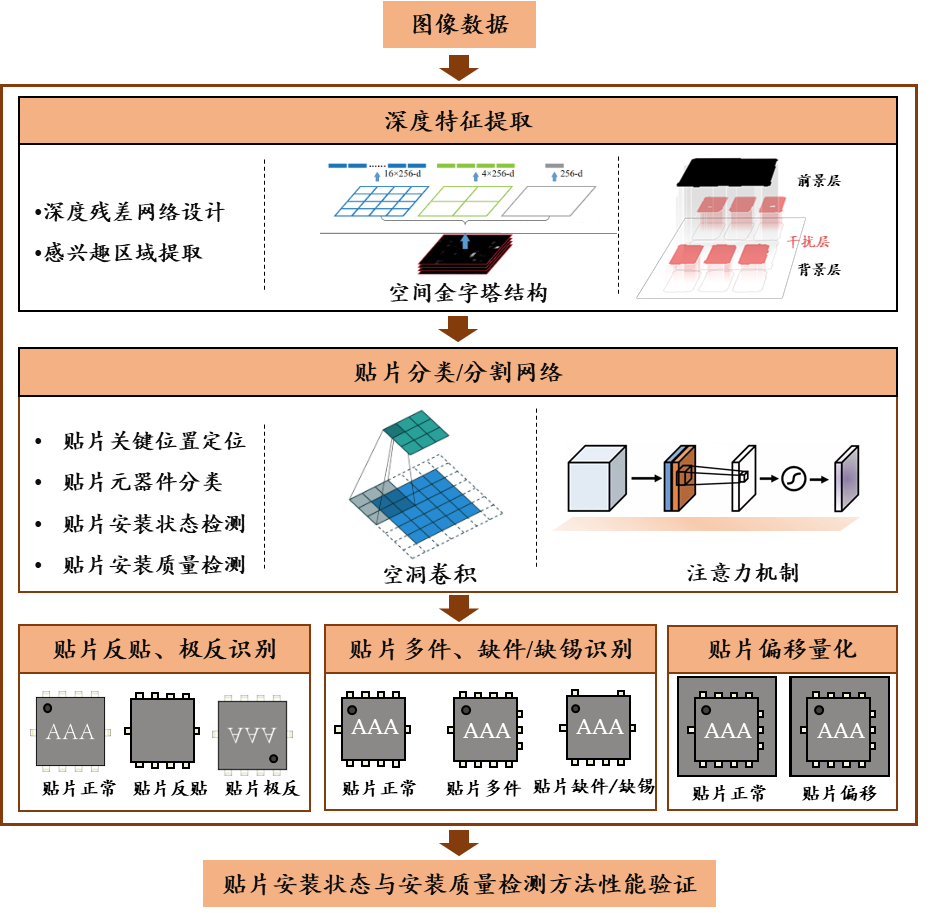


图6 贴片安装状态和安装质量检测方法

**② 集成电路贴片标识自适应高速辨识方法研究**：贴片标识识别的难点在于数据模糊及照度较低、外观多样性较高，前者可通过低照度图像感知质量提升方法进行改善，因此贴片标识识别的关键在于外观多样性较高的问题。通过设计基于残差结构的方向纠正网络，完成标识方向识别和修正；融合弱监督学习理念，设计动态标签更新策略和大间距Softmax函数，实现标识字符自适应识别任务。

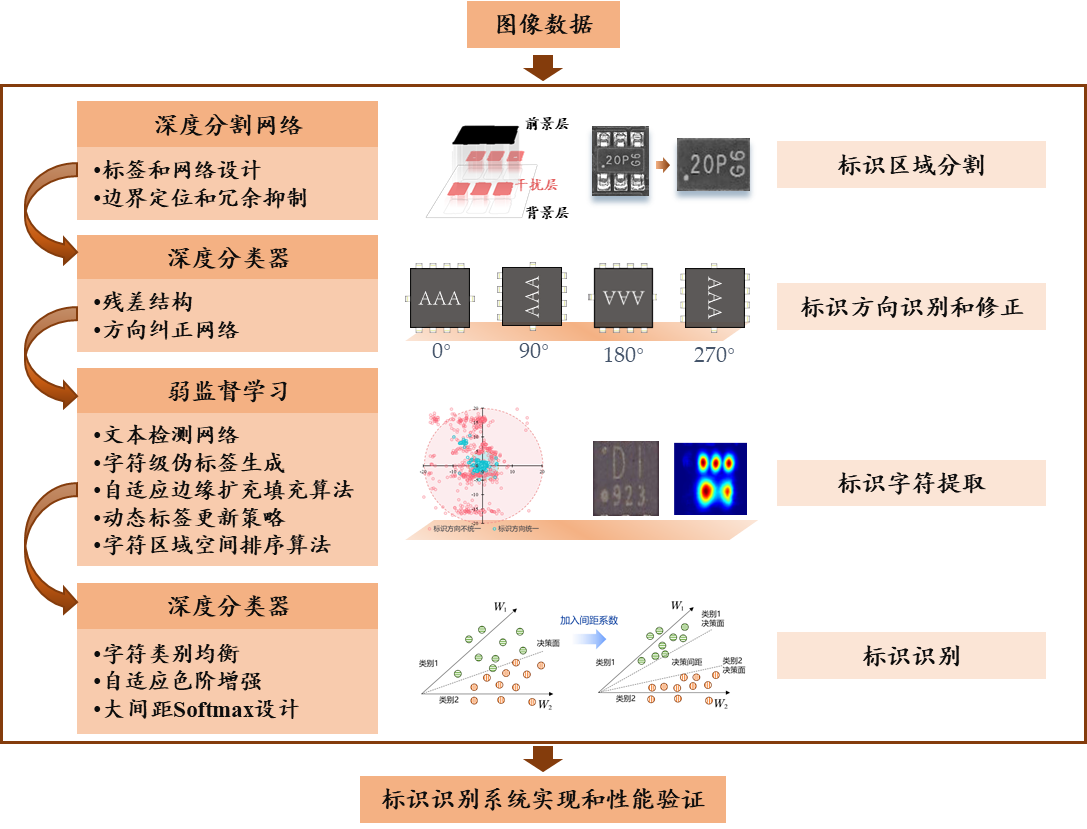
****

图7 贴片标识自适应高速辨识技术研究

**③ 基于在线持续学习的模型参数自适应更新策略研究：**面向动态环境下流式数据中产生的新缺陷样本数据，确保检测模型参数自适应更新是适配实际生产环节的核心任务。本项目拟采用基于回放的方法设计内存缓存区，从而应对灾难性的遗忘问题；研究基于对抗性思维的自适应在线持续学习，通过映射高维特征空间来完成新数据和旧数据极大极小博弈，混淆判别器以对齐新数据和旧数据特征分布，从而实现动态调整检测模型的参数。

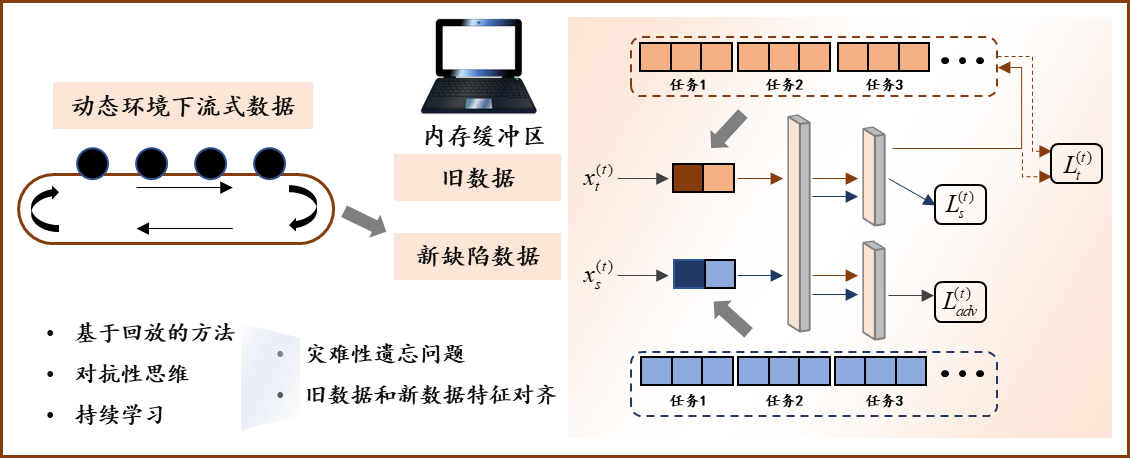


图8 基于在线持续学习的模型参数自适应更新策略研究

最后是对上述三个研究内容的综合应用。本项目拟通过构建多边协同的集成电路贴片工艺智能化质量检测平台来解决集成电路工程领域内所存在的一些瓶颈问题。通过运用算法思维与软件工程思想从实际应用角度完成平台系统的集成实现和性能验证，最后基于平台系统完成流水线作业方式下集成电路贴片在线智能检测任务。需要指出，上述研究内容并不是各自隔离的，而是相互关联、彼此支撑、从而共同组成一个有机的整体。通过智能化检测平台，将质量检测技术与信息技术紧密结合，深入挖掘对提高产品质量有价值的数据，服务于质量管理与质量控制，从而提高企业的生产决策支持服务能力与产品质量追溯管理能力，更好地响应国家制造强国的战略目标。

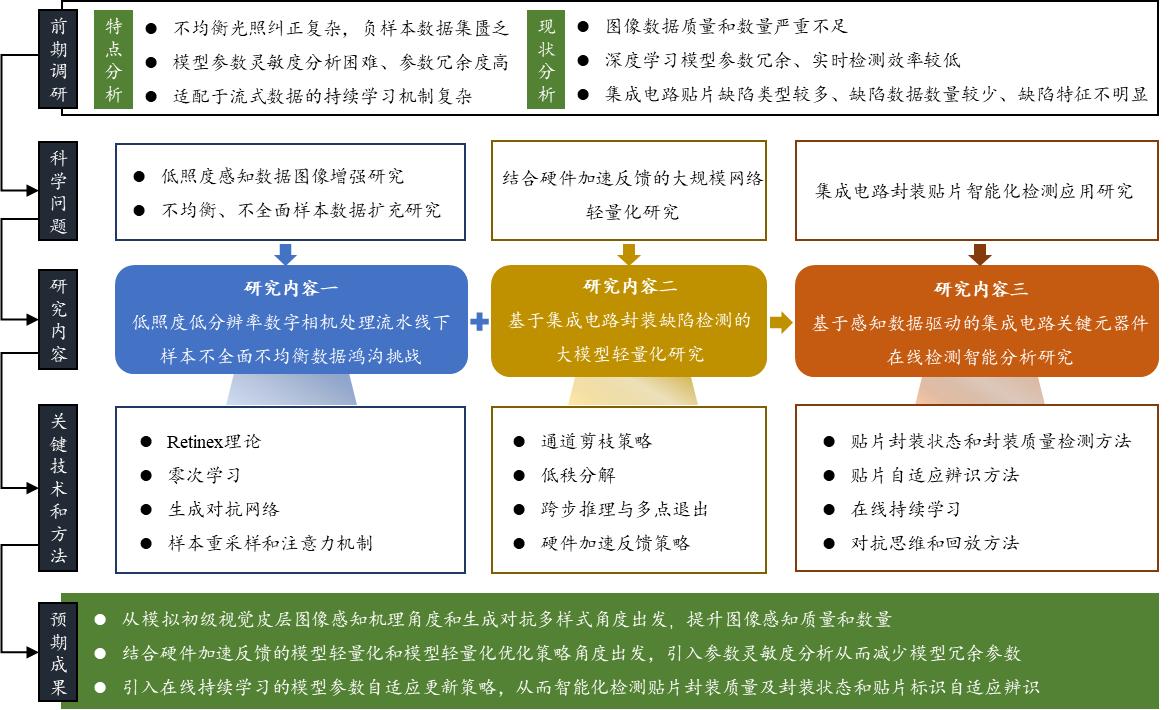


图9 本项目总体研究方案