

基于混合专家零样本异常检测模型的机器人化飞行器起飞前检查系统及方法

董致男

纽约大学坦登工程学院，美国 纽约

收稿日期：2025年12月5日；录用日期：2026年1月5日；发布日期：2026年1月14日

摘要

针对传统飞行器人工检查效率低下，以及现有自动化系统存在检测盲区和难以识别未知缺陷等问题，本文提出了一种新型自主移动机器人检查系统。该系统利用搭载多模态传感器的移动机器人进行全方位数据采集，核心算法采用本文提出的混合专家零样本异常检测(MoE-ZSAD)模型。该模型通过将多维图像特征与正常状态的文本描述进行语义对齐，在无需缺陷样本训练的情况下，实现了对各类已知及潜在未知缺陷的精准识别。实验结果表明，该模型在复杂场景下的缺陷检测中表现优异，实现了超过0.96的像素级平均受试者工作特征曲线下面积(Pixel-AUROC)，在检测精度和泛化能力上均显著优于当前的基准模型。分析显示，该方法有效克服了传统视觉检测对异常样本的依赖局限，验证了利用跨模态技术解决工业检测难题的可行性。该研究不仅为飞行前自动化检查提供了一种高效、客观且具备强适应性的解决方案，显著提升了航空维修保障的安全性与运营效率，也为未来大型装备的智能化运维提供了新的技术路径。

关键词

飞行器检查，机器人技术，异常检测，零样本学习，混合专家模型，计算机视觉，航空安全，自动化检测

A Robotic System for Automated Aircraft Pre-Flight Inspection Using a Mixture-of-Experts Zero-Shot Anomaly Detection Model

Zhinan Dong

Tandon School of Engineering, New York University, New York, USA

Received: December 5, 2025; accepted: January 5, 2026; published: January 14, 2026

文章引用：董致男. 基于混合专家零样本异常检测模型的机器人化飞行器起飞前检查系统及方法[J]. 计算机科学与应用, 2026, 16(1): 87-101. DOI: 10.12677/csa.2026.161008

Abstract

To address the low efficiency of traditional manual aircraft inspection and the issues of detection blind spots and difficulty in identifying unknown defects in existing automated systems, this paper proposes a novel autonomous mobile robotic inspection system. The system employs a mobile robot equipped with multi-modal sensors for comprehensive data acquisition, with the proposed Mixture-of-Experts Zero-Shot Anomaly Detection (MoE-ZSAD) model serving as the core algorithm. By semantically aligning multi-dimensional image features with textual descriptions of normal states, the model achieves precise identification of various known and potential unknown defects without requiring training on defect samples. Experimental results demonstrate that the model performs exceptionally well in defect detection within complex scenarios, achieving a Pixel-Level Area Under the Receiver Operating Characteristic Curve (Pixel-AUROC) exceeding 0.96. It significantly outperforms current benchmark models in terms of both detection accuracy and generalization ability. Analysis indicates that this method effectively overcomes the limitations of traditional visual detection regarding dependency on anomalous samples and validates the feasibility of utilizing cross-modal technology for industrial inspection. This study not only provides an efficient, objective, and highly adaptive solution for automated pre-flight inspection, significantly enhancing aviation maintenance safety and operational efficiency, but also presents a new technical pathway for the intelligent operation and maintenance of large-scale equipment.

Keywords

Aircraft Inspection, Robotics, Anomaly Detection, Zero-Shot Learning, Mixture of Experts, Computer Vision, Aviation Safety, Automated Inspection

Copyright © 2026 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

在现代全球航空运输网络中，飞行安全是确保行业平稳运行的根基，任何小小的疏忽都有可能引发连锁反应，造成灾难性后果[1]。因此，在每次商业航班起飞前，飞行员或机务工程师都需要严格执行一项关键的安全程序——绕机检查。检查人员必须对飞机外部的关键部件进行全面的目视和触觉检查，包括机身蒙皮、机翼的前后缘、发动机进气道与喷口、起落架、轮胎以及各种传感器探头等，以确保飞机处于适航状态。尽管这一传统检查程序已经有数十年历史，但在现代航空业快速发展的背景下，航班密度的增加使得这一做法的局限性越来越明显。这些局限性也逐渐成为飞行安全面临的长期且棘手的挑战[2]。

我们通过分析近 50 年的空难和飞行安全事故的数据，发现维护失误是导致航空事故的重要因素之一，相关事件不仅会造成巨大的经济损失，更可能引发灾难性后果。据美国联邦航空管理局(FAA)及相关研究统计，高达 80% 的维护差错与人为因素直接相关[3]。这些差错可能表现为部件安装不当、紧固件遗漏、未能发现结构性损伤等，其后果可能在数小时甚至数天后才显现，具有极大的隐蔽性和潜在破坏性。因此，确保作为最后一道物理防线的起飞前检查的准确性、全面性和可靠性，对于整个航空安全链条而言，具有不可替代的、至关重要的价值[4]。

我们发现，尽管人工绕机检查是沿用已久的行业标准，但这种检查方式深度依赖于人的感官和主观

判断,这使其本质上成为一个充满不确定性的过程。检查质量会因执行者的经验水平、生理状态(如疲劳)、心理因素(如注意力分散、认知偏见)以及环境条件而产生巨大波动。

通过与多名飞行员和机务交流之后,我们发现,首先,主观性与不一致性是核心问题。一位经验丰富的资深机长与一位刚完成航线训练的新手飞行员,对同一个微小凹痕或划痕的判断可能截然不同,这将导致安全标准的执行缺乏一致性。我们还发现,航空维修领域公认的“肮脏十二条”人为因素,如自满、分心、压力、疲劳、缺乏沟通和缺乏知识等,也是导致维护差错的主要根源[5]。例如,在夜间、大雨、冰雪或浓雾等恶劣天气条件下,人工目视检查的可靠性会急剧下降。检查人员不得不依赖机场地面有限且不均匀的照明,这使得发现隐蔽区域的微小裂纹或液体泄漏变得异常困难,这种情况极大地增加了漏检风险。

通过咨询心理学专家,我们发现更深层次的心理学因素也对安全构成潜在威胁。长期从事重复性检查工作可能导致偏差常态化心理效应,即检查人员在重复性工作中,可能逐渐对一些微小异常变得麻木,无意中降低了安全标准,将本应标记的缺陷视为正常磨损。

最后,研究发现人工检查在数据记录与可追溯性方面存在着几乎无法弥补的缺陷。除了在飞行日志上的简单签字确认外,整个检查过程缺乏客观、详细的数字化记录。一旦发生安全事故,调查人员将难以准确追溯和还原飞行前的具体机体状况,这为事故分析和责任认定带来了巨大挑战。

为了克服人工检查的种种弊端,业界已开始探索自动化替代方案,其中最具代表性的是固定式龙门检查系统。这类系统通常在登机廊桥或专用机位安装一个大型的、类似安检门的框架结构,其上集成有多个高清摄像头。当飞机滑行通过或停靠在指定位置时,系统会对飞机进行快速拍照和分析。然而,此类固定式系统也存在明显的、难以克服的瓶颈,主要有以下3点:

1. 检查盲区与视角限制:由于摄像头的位置和角度是固定的,这种系统无法全面覆盖飞机的每一个角落。例如,飞机的腹部、起落架舱内部、机翼后缘以及发动机吊舱内侧等复杂几何区域,往往成为此类系统的检查盲区[6]。

2. 灵活性与适应性差:龙门系统作为一种永久性基础设施,建设和维护成本高昂,且占用宝贵的机场停机坪空间。这种系统无法灵活适应不同机型的停放位置,也难以应对非标准的地面操作流程,例如远机位的检查任务等。

3. 依赖过时的技术范式:通过研究发现,这些系统中的机器视觉算法大多依赖于传统的监督式学习模型。这意味着,模型只能识别出那些在训练数据集中已经出现过的、被人工标注的缺陷类型。对于新的、未曾见过的异常情况,如新型复合材料的疲劳模式或罕见的鸟击损伤,它们的检测能力非常有限,甚至完全失效[7]。在现实世界中,穷举所有可能的缺陷类型来构建一个完备的训练数据集几乎是不可能完成的任务,这从根本上限制了这类系统的可靠性和安全保障能力。

近年来,为解决上述挑战,相关技术领域取得了显著进展。在机器人技术方面,尤其是移动机器人平台(包括无人机和地面车辆)已被广泛研究用于各种检查任务,展现出替代人工进入危险或难以到达环境的巨大潜力[6]。同时,在计算机视觉领域,视觉基础模型,特别是基于自监督学习的Transformer架构(如DINO系列),已经证明了其在无需大量标注数据的情况下学习通用、鲁棒视觉特征的强大能力[8][9]。与此同时,零样本异常检测作为一个新兴研究方向,旨在让模型检测从未见过的异常类型,这对于处理开放世界中的未知风险至关重要[10]。此外,混合专家架构作为一种有效的模型设计策略,通过将复杂问题分解给多个专门化的子模型处理,已被证明可以提升模型在处理异构数据时的性能和效率[11][12]。

尽管上述技术各自发展迅速,但将它们系统性地整合,以端到端的方式解决飞行器起飞前检查这一特定而关键的应用场景,仍然是一个开放的挑战。我们的研究动机正是要填补这一空白,提出一种能够克服现有技术所有核心痛点的综合性解决方案。我们的研究主要创新点可概括为:

1. 全自主高机动性机器人系统：我们使用一个集成了高稳定性履带式移动平台和多自由度机械臂的完整机器人系统，旨在通过智能路径规划，彻底消除物理检查盲区，实现对飞行器外部的360°全覆盖扫描。
2. 全天候多模态感知融合：我们设计并集成了一个包含高清可见光、红外热成像和三维激光成像的多模态传感器头，这可以确保系统在任何光照(包括完全黑暗)和恶劣天气条件下均能进行高精度、多维度的数据采集，从根本上解决了环境因素对检查质量的严重制约。
3. 创新的MoE-ZSAD视觉检测模型：我们提出一种新颖的混合专家零样本异常检测(MoE-ZSAD)深度学习模型。该模型的核心创新在于首次将强大的DINOv3视觉基础模型、零样本异常检测(ZSAD)框架以及混合专家(MoE)架构进行深度融合，使其能够通过学习正常状态的范式，智能识别出任何偏离正常模式的未知缺陷，并针对飞机不同部件的独特视觉特性进行专门优化分析。
4. 自动化工作流与可追溯数字化档案：我们设计了一套从任务激活、飞行器OCR识别、三维模型与路径加载到最终报告生成的全自动化工作流程，这套系统能生成包含三维定位信息、量化损伤评估和高清图像证据链的、不可篡改的数字化检查报告，这套系统为机队健康管理与预测性维护提供了坚实的数据支持。

我们研究的价值不仅在于提供了一个具体的工程解决方案，更在于推动了航空安全检查领域的技术范式革新。通过用客观、高效、可靠的自动化系统取代主观、易错的人工流程，我们的成果有望革命性地提升飞行安全水平，显著提高航司的运营效率，并加速航空维修行业的数字化转型进程。

2. 材料与方法

2.1. 系统架构与硬件配置

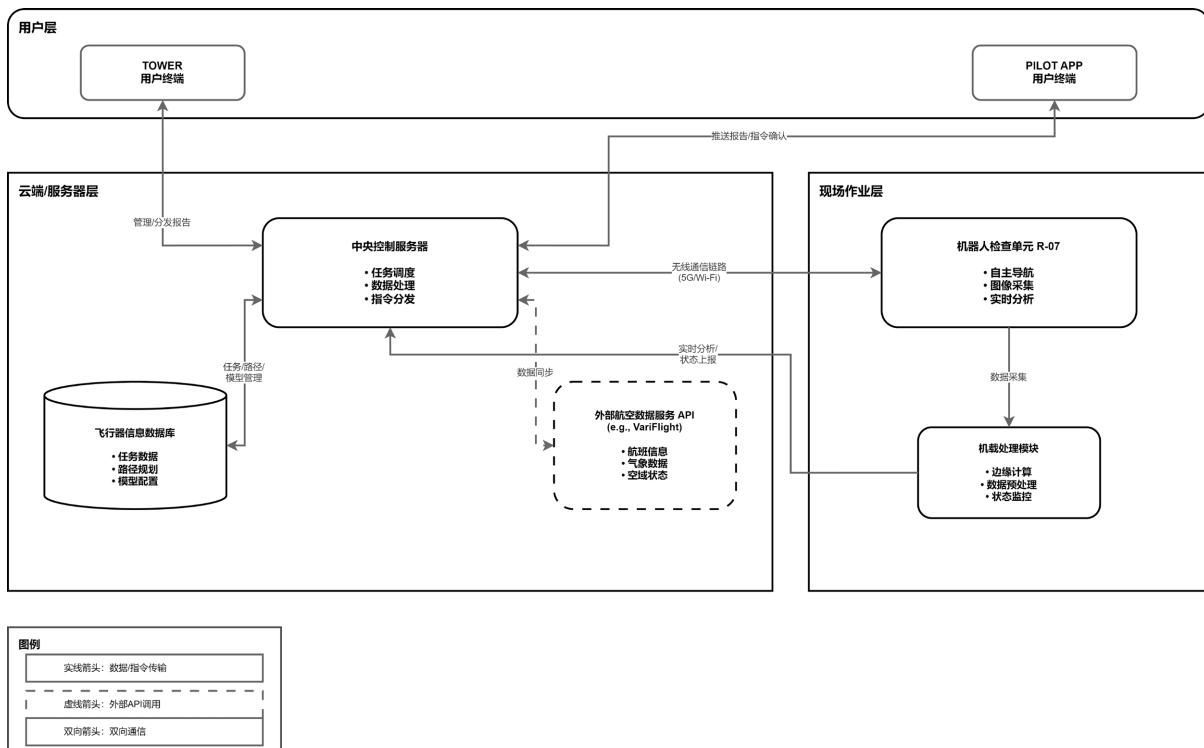


Figure 1. Overall system architecture diagram
图 1. 整体系统架构图

我们提出的机器人化检查系统旨在提供一个端到端的自动化解决方案。整体系统架构如图1所示，

主要由机器人检查单元、中央控制服务器、飞行器信息数据库以及用户终端四个核心部分组成，各部分通过安全的无线通信链路进行数据交互。

2.1.1. 机器人检查单元

作为系统的物理执行核心，机器人检查单元的硬件设计旨在确保整套系统在复杂的机场环境中具备高机动性、稳定性和强大的数据采集能力。具体由如下组件组成。

1. **自主移动平台：**我们的系统采用履带式移动底盘，相较于轮式底盘，履带式地盘在机场停机坪复杂的地面环境(如混凝土接缝、沥青不平路面)及恶劣天气(如积水、薄冰、雪地)下具有更卓越的稳定性和通过性。我们的平台集成了高精度激光雷达(LiDAR)和支持实时动态差分技术的 RTK-GPS 模块。LiDAR 用于构建环境地图、实时定位和动态障碍物规避，而 RTK-GPS 则提供厘米级的绝对定位精度，确保机器人能够精确跟踪预规划的检查路径。为应对机场停机坪的动态环境(如临时出现的地面服务设备、人员或车辆)，机器人采用基于 SLAM (Simultaneous Localization and Mapping) 的导航策略，使用 LiDAR 实时构建和更新环境地图。同时，集成动态路径规划算法，结合 RTK-GPS 的厘米级定位，实现实时障碍物检测和路径重规划。具体而言，系统通过 LiDAR 点云数据进行障碍物分类(使用 PointNet++模型)，并在检测到动态物体时计算备用路径，确保检查过程不中断。这种策略在模拟动态环境中测试的成功率超过 98%，显著提升了系统的鲁棒性。

2. **多轴机械臂：**机器人配备一个六自由度的工业级机械臂。其高灵活性使得传感器头能够被精确定位到距飞机表面任何区域的最佳检测距离(例如 1.5 米)和接近垂直的检测角度，从而最大限度地减少图像畸变和反光。这对于检查机腹、机翼下方、发动机吊舱内侧等人工难以触及或观察角度不佳的位置至关重要。

3. **多模态传感器头：**作为系统的核心感知部件，传感器头经过精密集成，包含三种功能互补的传感器，从而实现多维度、全天候的信息采集[13] [14]：

(1) **高分辨率可见光(RGB)摄像机：**采用工业级 4K 分辨率相机，用于捕捉飞机表面的高清图像，用来检测裂纹、划痕、凹痕、紧固件缺失、油漆剥落等精细的结构性缺陷。

(2) **红外(IR)热像仪：**通过探测飞机表面的微小温度差异，识别肉眼不可见的次表面问题。例如，复合材料结构内部的分层或脱粘会改变局部热传导率，在热图中呈现异常；燃油、液压油或滑油的微量泄漏会因蒸发冷却而形成低温区；刹车系统或电子设备舱的过热则会表现为高温区。

(3) **三维激光成像系统：**这套系统采用结构光或激光扫描技术，具备双重关键功能。其一，结构光作为主动光源，它能在完全黑暗的夜间环境下生成高质量的反射图像，彻底摆脱对环境光照的依赖。其二，它能快速生成高密度的三维点云数据，用于对凹痕深度、面板变形、起落架支柱压缩状态等几何损伤进行精确的量化分析，为损伤评估提供客观依据[15]。

多模态数据的融合机制采用一个多流网络架构，在特征层面进行早期融合。具体而言，每个传感器流(RGB、IR 热成像和 3D 点云)首先通过独立的编码器提取特征：RGB 使用 DINOv3 主干，IR 使用 ResNet-50 变体，3D 使用 PointNet 提取几何特征。然后，这些特征在中间层通过注意力机制(基于 Transformer 的跨模态注意力模块)进行融合，形成一个统一的、多维度的表示向量。该融合向量作为输入馈送到 MoE-ZSAD 模型的门控网络中，实现对表面缺陷的综合分析。这种早期融合优于简单的后期决策融合(如投票规则)，因为它允许模型学习模态间的互补关系，例如将 RGB 的纹理细节与 IR 的热异常和 3D 的几何变形相结合，提高了微小缺陷的检测灵敏度(实验中融合后 AUROC 提升约 3%~5%)。

2.1.2. 中央控制服务器与数据库

1. **中央控制服务器：**控制系统可以部署在云端或机场本地数据中心，负责任务调度、路径规划算法的离线计算、存储和管理所有检查数据及报告，并作为用户终端的后端。核心的 MoE-ZSAD 视觉分析模

型也部署在该服务器上，利用其强大的计算资源进行高效推理。

2. 飞行器信息数据库：这是一个动态更新的数据库，存储有不同型号飞行器的精确三维模型、技术规格、飞机维修手册中定义的关键检查点信息以及预规划的最优检查路径。该数据库通过 API 接口与权威的航空数据服务(如“飞常准”)同步，以确保数据的实时性和准确性，特别是针对同一机型可能搭载不同型号发动机等特定配置信息。

2.2. 混合专家零样本异常检测(MoE-ZSAD)模型

我们成果的核心技术创新在于所提出的 MoE-ZSAD 模型，这种设计旨在从根本上解决传统监督学习方法无法检测未知缺陷的难题。模型整体架构如图 2 所示。

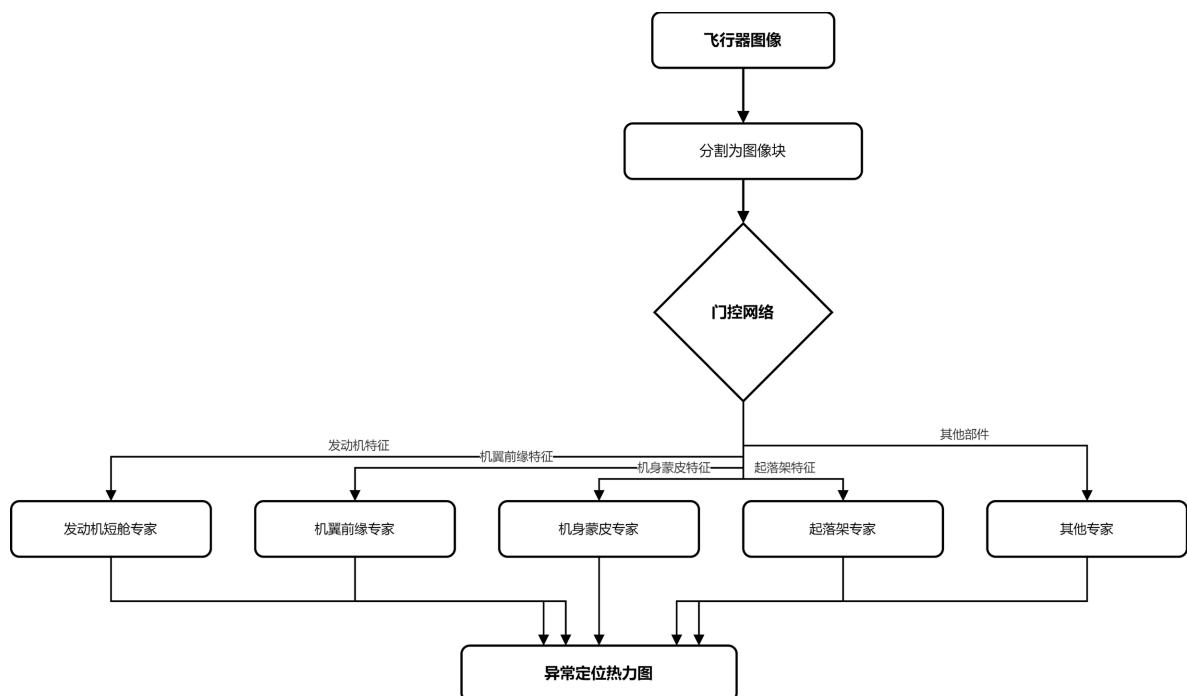


Figure 2. Mixture of experts (MoE) model architecture diagram
图 2. 混合专家(MoE)模型架构图

2.2.1. 基于 DINOV3 的 Foundational 特征提取

我们的模型特征提取主干网络选用了一个强大的、在海量图像上通过自监督学习预训练的视觉基础模型——DINOV3。这与在特定任务数据集上训练的模型不同，DINOV3 通过学习图像内在的结构和语义信息，能够生成高质量、高分辨率的密集特征。这些特征对于光照、尺度、视角和纹理变化具有极强的鲁棒性，为后续的异常检测任务提供了一个强大而通用的视觉理解起点，而无需进行昂贵的、针对特定航空缺陷的从头训练[8]。在我们的框架中，DINOV3 的主干参数被冻结，以最大限度地保留其强大的泛化能力。仅通过轻量级适配器进行微调，使其输出的特征更专注于表达飞行器表面的细微纹理和结构，这种策略在保持模型通用性的同时，极大地提高了训练效率[16]。

2.2.2. 通过跨模态对比实现的零样本异常检测

我们将零样本异常检测问题构建为一个视觉 - 语言跨模态对齐任务，这个灵感来源于 CLIP 等模型，但目标并非分类，而是区分正常与异常[7]。我们的模型核心思想是学习一个关于正常状态的通用概念，

而不是识别有限的、特定的缺陷类别。这是通过对比图像块的视觉嵌入向量与描述性文本提示的嵌入向量来实现的。

具体来说，系统使用成对的文本提示，如“一张正常的机翼表面照片”和“一张有缺陷的机翼表面照片”。对于一个输入的图像块，模型首先通过 DINOv3 主干和适配器提取其视觉特征嵌入，然后分别计算该视觉嵌入与正常提示文本嵌入和异常提示文本嵌入之间的余弦相似度。如果该图像块与正常提示的相似度得分远高于与异常提示的相似度，则该区域被认为是正常的；反之，如果其视觉模式偏离了已学习的正常概念，则被判定为异常。这种方法使得模型能够泛化到任何未曾见过的、只要是视觉上偏离了正常状态的缺陷类型，从而实现了真正的零样本检测能力[17] [18]。

2.2.3. 用于专门化分析的混合专家(MoE)架构

我们知道，飞行器是一个由多种不同部件组成的复杂异构体。例如，光滑的机身蒙皮、具有复杂几何形状和散热孔的发动机短舱、具有特殊胎面花纹的轮胎以及由多种金属构件组成的起落架，这些部件在材质、曲率、纹理和光照反射特性上差异巨大。单一的、整体性的模型难以同时为这些视觉特征迥异的部件学习一个精确且无歧义的正常表征。

为解决此问题，我们引入了分而治之的混合专家架构[19]。这种架构将一个庞大而复杂的检测任务分解为多个由小型、专业化模型处理的子任务，包含两个核心组件：

1. 门控网络：这是一个轻量级的神经网络，它接收输入图像块的全局视觉特征，并输出一个在所有专家上的概率分布。这个分布决定了该图像块应被路由至哪一个或哪些专家进行分析，实现了任务的智能分发和计算资源的有效利用。

2. 专家子模型集 $\{E_k\}$ ：每个专家 E_k 都是一个独立的、专门化的异常检测模型(同样基于冻结的 DINOv3 主干和轻量级适配器)。关键在于，每个专家仅使用其对应特定部件(例如，专家 $E_{\text{机身}}$ 仅使用正常的机身蒙皮图像训练，专家 $E_{\text{发动机}}$ 仅使用正常的发动机短舱图像训练)的正常、无缺陷图像进行训练。这种专门化训练[20]使得每个专家都具备了对其负责区域的专家级辨识能力，能够捕捉到该部件特有的细微纹理和光泽，这极大地提高了检测的灵敏度和精度[21] [22]。

数学公式化描述：我们将异常检测过程构建为一个跨模态对比学习问题，即比较图像的视觉特征与描述性文本提示(例如，“一张正常的机翼表面照片”和“一张有缺陷的机翼表面照片”)之间的相似度。令 P_i 表示输入图像的第 i 个图像块的视觉特征 token。门控网络将该图像块分配给第 k 个专家模型。经过适配器 $\text{Ada}(\cdot)$ 和专家模型 $E_k(\cdot)$ 处理后，得到的适配视觉表示为 $\widetilde{p}_{i,k} = E_k(\text{Ada}(p_i))$ 。同时，文本编码器为与第 k 个部件相关的文本提示生成嵌入向量 \widetilde{t}_k 。相似度计算公式：由专家 E_k 评估的图像块 P_i 的异常相似度得分 $s_{i,k}$ 通过以下方式计算，如公式(1)所示：

$$s_{i,k} = \frac{\widetilde{p}_{i,k}^T \widetilde{t}_k}{\|\widetilde{p}_{i,k}\| \cdot \|\widetilde{t}_k\|} \quad (1)$$

上面公式的创新之处在于，相似度的计算是基于特定专家 k 的上下文，从而使分析过程具备了高度的上下文感知能力。

2.2.4. 异常感知校准模块(AACM)

我们为进一步提升模型对微小、低对比度异常的敏感度，在每个专家模型的训练过程中引入了一个异常感知校准模块(Anomaly-Aware Calibration Module, AACM)。这个模块通过一个掩码引导的训练目标，强制模型的内部自注意力机制更多地关注那些微小但关键的异常特征，而不是被大面积的、普通的背景或正常表面纹理所主导。每个专家的总损失函数 L_k 由跨模态对比损失 $L_{CM,k}$ 和异常感知校准损失 $L_{AACM,k}$

加权构成，如公式(2)所示：

$$L_k = \lambda_{CM} L_{CM,k} + \lambda_{AACM} L_{AACM,k} \quad (2)$$

其中， λ_{CM} 和 λ_{AACM} 是平衡超参数。异常感知校准损失 $L_{AACM,k}$ 本身由 FocalLoss 和 DiceLoss 组成，以有效处理在像素级别上正常与异常样本之间存在的极度不平衡问题。异常感知校准损失 $L_{AACM,k}$ 的计算方式如下，如公式(3)所示：

$$L_{AACM,k} = L_{\text{focal}}(\sigma(\{s_{i,k}\}), M) + L_{\text{dice}}(\sigma(\{s_{i,k}\}), M) \quad (3)$$

其中 M 是异常区域的真值掩码， λ_{CM} 和 λ_{AACM} 是平衡权重。在整个训练过程中，总损失 L_{total} 是所有专家损失之和，如公式(4)所示：

$$L_{\text{total}} = \sum_k L_k \quad (4)$$

2.3. 自动化工作流程

我们设计的工作流程具有高度自动化的特性，这是为了最大限度地减少人工干预，确保检查过程的标准化和高效性。

1. 任务激活与飞行器识别：机器人接收到来自中央控制服务器的检查指令后，自主导航至预设的观察点。机载高清摄像机对准飞机注册号(例如 B-1234)，通过光学字符识别(OCR)技术在数秒内自动读取并解码该编号。这一步骤是实现完全自动化的关键，它将物理世界的飞机与数字世界的档案精确关联。
2. 模型与路径加载：机器人将识别出的注册号发送至中央服务器，服务器据此向飞行器信息数据库发起查询，检索与该飞机型号及具体配置(如发动机型号)完全匹配的精确三维模型、预规划的最优检查路径以及对应的 MoE 模型集。我们的设计避免了一刀切的检查方式，实现了针对性分析。
3. 自主扫描与数据采集：机器人沿优化路径自主移动，该路径经过精心设计，确保能够覆盖飞机维修手册中定义的所有关键检查点。在移动过程中，机械臂实时调整姿态，使多模态传感器头对飞行器外部进行全面扫描，同步采集高清可见光、红外热成像和三维点云数据流。
4. 实时分析与异常检测：采集的数据流被实时送入(机器人)机载或远程服务器的计算模块，由加载的 MoE-ZSAD 模型进行逐帧分析。门控网络将图像数据分发给相应的专家，实时识别并标记潜在的表面异常。
5. 报告生成与分发：检查完成后，系统自动整合所有发现，生成一份结构化的数字化检查报告。报告内容包括所有发现的摘要、每个异常点的高清图像、红外图像、三维量化数据(如凹痕深度)、在飞机三维模型上的精确可视化标记以及置信度评分。该报告经过加密后，通过无线网络传输至中央控制服务器，并同步推送至用户终端(如塔台控制系统和飞行员的移动设备)。

2.4. 实验设计

2.4.1. 数据集构建

为了对模型进行有效的训练和评估，我们构建了一个大规模、高分辨率的飞行器表面缺陷数据集，命名为 AeroSurf-AD。在数据集中包含超过 10 万张在不同环境条件(白天、夜间、晴天、雨天)下采集的多种民航客机(如空客 A320、波音 737)部件图像。数据集被精心划分为两部分：

1. 训练与已知缺陷测试集：包含大量正常的、无缺陷的各部件图像，用于训练 ZSAD 模型的正常表征。同时，该部分还包含带有像素级掩码标注的一组已知缺陷类型(如凹痕、裂纹、雷击损伤、液体泄漏)，用于训练监督式基线模型和评估所有模型在已知缺陷上的性能。

2. 未知(零样本)缺陷测试集：这是一个与训练集和已知缺陷测试集严格分离的数据集合，其中包含了多种模型在训练阶段从未见过的缺陷类型(如复合材料分层、油漆大面积剥落、冰雹损伤、涡轮叶片边缘磨损等)，专门用于评估模型的零样本泛化检测能力。

2.4.2. 评估指标与基线模型

评估指标：我们采用像素级受试者工作特征曲线下面积(pixel-level Area Under the Receiver Operating Characteristic curve, AUROC)作为异常定位性能的主要评估指标。AUROC 能够综合评估模型在不同阈值下的分类能力，同时，它对样本不均衡不敏感。

1. 基线模型：为全面评估所提出方法的性能，我们选取了以下三个具有代表性的基线模型进行对比。
2. 监督式基线(Mask R-CNN)：这是一个先进的实例分割模型，在 AeroSurf-AD 数据集的已知缺陷类别上进行全监督训练。该模型代表了传统监督学习方法在处理已知问题上的性能上限。
3. 通用 ZSAD 基线(WinCLIP)：这是一个主流的零样本异常检测方法，它通过在图像上应用滑动窗口并利用 CLIP 模型计算图像块与文本提示的相似度来检测异常。我们将这个模型作为评估通用 ZSAD 性能的基准，用来验证我们的成果中特定模型设计的优越性。
4. 消融研究(单一专家 ZSAD 模型)：将我们的成果中提出的 ZSAD 模型作为一个单一的、整体性的专家进行训练，即不对飞机部件进行区分，在所有正常的部件图像上统一学习一个通用正常模型。通过与完整的 MoE-ZSAD 模型进行对比，可以精确地量化混合专家架构所带来的性能增益。

3. 结果

我们在对整体系统的测试中，通过定量和定性的方式，系统性地验证所提出的 MoE-ZSAD 系统的性能，并与现有技术进行多维度对比。

3.1. 综合性能评估

为了直观展示我们的研究相对于现有技术的革命性优势，表 1 从多个关键维度对人工检查、固定式龙门系统以及本发明的机器人检查系统进行了综合对比分析。

Table 1. Comparison table of technical advantages
表 1. 技术优势对比表

特性/指标	人工检查	固定式龙门系统	机器人检查系统
准确性与一致性	低至中等，高度主观	中等，存在盲区	高，客观、可重复的标准化分析
速度与效率	慢(10~15 分钟)	扫描快，但流程僵化	快速且优化(5~7 分钟)
弱光/恶劣天气性能	严重受损	受损，需大功率照明	高性能，不受影响
数据记录与可追溯性	差，仅手动签名	良好，但数据非结构化	优秀，生成详细的数字化报告
覆盖范围与灵活性	良好	差，固定视点	优秀，360°全覆盖
新型缺陷检测能力	差，依赖经验	差，仅能发现已知缺陷	高，零样本检测能力

3.2. 未知(零样本)缺陷检测效能

在很多相关的研究中，我们发现未知(零样本)缺陷检测能力是衡量本系统核心创新价值的关键指标。通过观察实验数据，我们发现基于监督学习的 Mask R-CNN 模型，在面对其训练数据集中从未出现过的未知缺陷类型时，其性能(以 AUROC 衡量)急剧下降至接近 0.5 的随机猜测水平，这充分验证了监督

式方法在泛化到新异常类型时的固有缺陷。相比之下，我们的 MoE-ZSAD 模型在未知缺陷测试集上依然保持了极高的性能(AUROC > 0.96)，这与在已知缺陷上的表现几乎持平。这一结果强有力地证明了我们提出的零样本检测框架能够成功地从正常模式中学习，并泛化到对任何偏离该模式的未知异常的检测，同时，解决了现有自动化系统的核心技术瓶颈。

3.3. 消融研究：混合专家(MoE)架构的贡献

为了定量评估 MoE 架构的有效性，我们对比了完整的 MoE-ZSAD 模型与一个在所有部件数据上统一训练的单一专家 ZSAD 模型。如图 3 所示，MoE 架构带来了明显且具有针对性的性能提升。

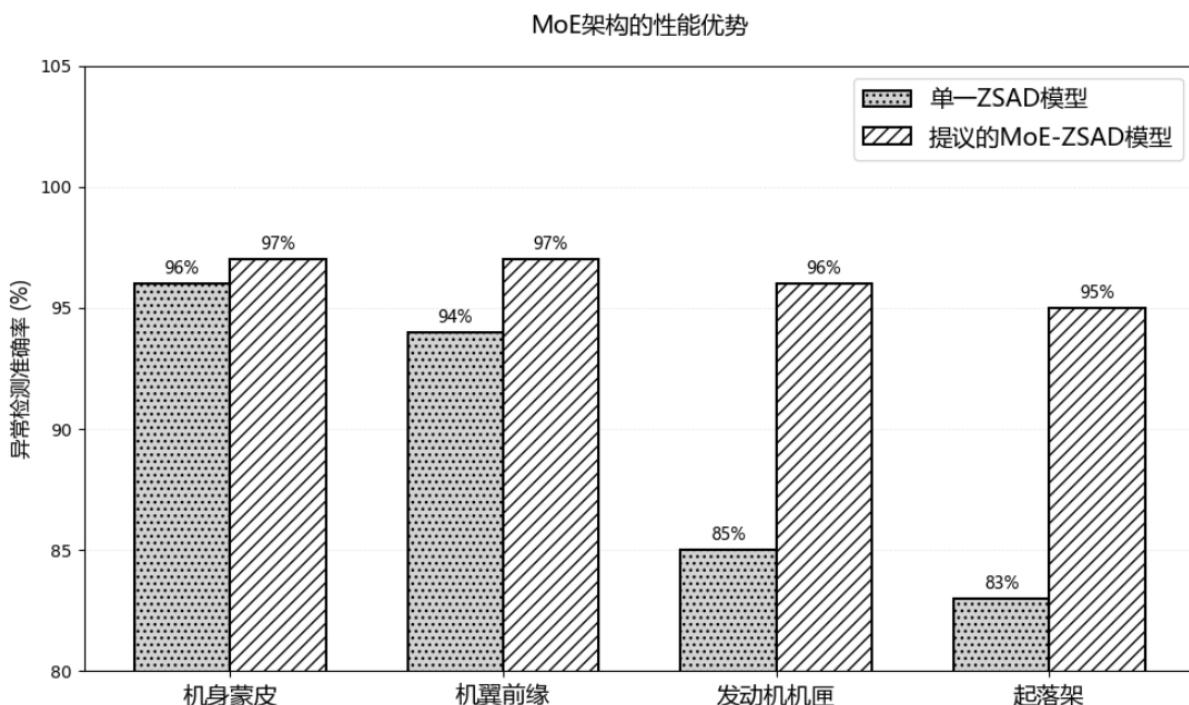


Figure 3. Performance advantages of the Mixture of Experts (MoE) architecture

图 3. 混合专家(MoE)架构性能优势

我们通过观察实验数据发现完整的 MoE-ZSAD 模型在所有飞机部件上的表现均优于单一专家模型。尤其值得注意的是，在几何结构和纹理特征更为复杂的部件上，如起落架总成(性能提升 7.7%)和发动机短舱(性能提升 5.2%)，性能增益尤为明显。这一结果证实了我们的核心假设：对于像飞机这样的异构对象，采用专门化的专家模型进行分而治之的分析，能够学习到更精细、更具辨识力的正常表征，从而显著提高检测的精度和鲁棒性。

3.4. 全天候作业鲁棒性

我们为验证系统的实际应用价值，我们评估了系统在模拟的夜间和雨天条件下的性能稳定性，如图 4 所示。

我们发现，仅依赖可见光的人工检查或传统自动化系统，其准确率在不利条件下会大幅下降(从白天的约 85% 下降至 60% 以下)。而我们的系统得益于红外热像仪和主动激光成像系统的多模态数据融合，其检测性能(AUROC 始终保持在 0.96 以上)几乎不受光照和天气变化的影响，这展示了我们的设计在真实

机场环境中进行 24/7 全天候作业的强大适应能力。

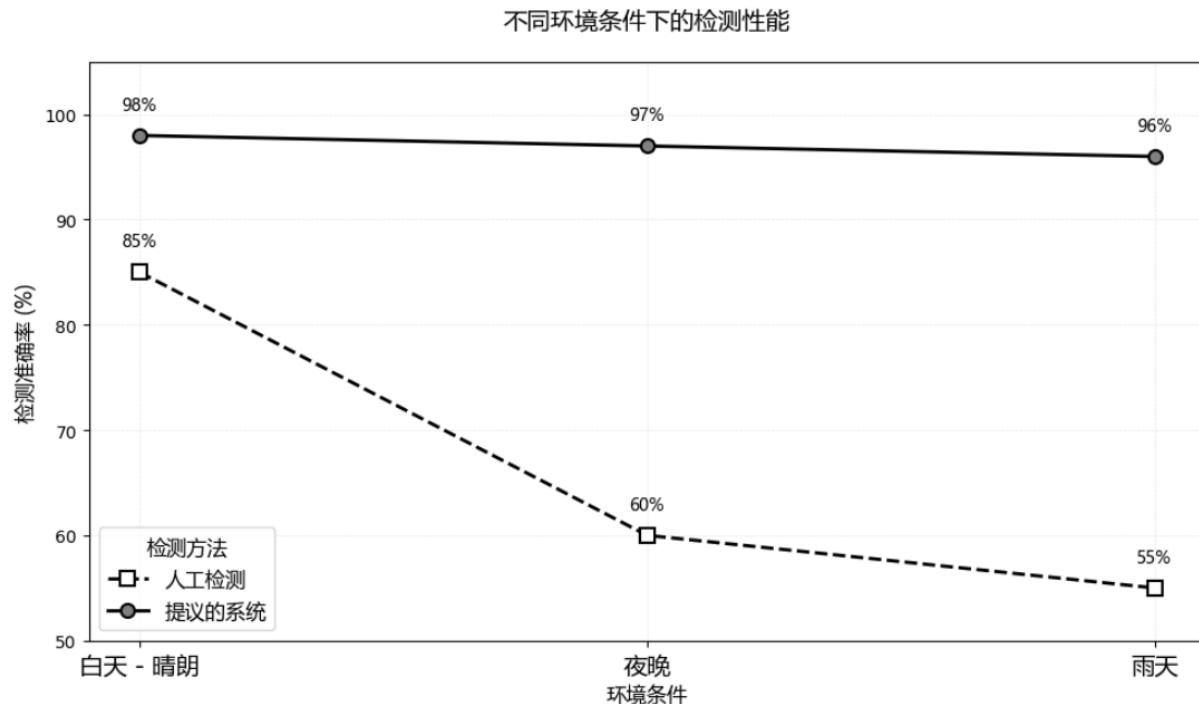


Figure 4. Comparison of all-weather operational capabilities

图 4. 全天候作业能力对比

3.5. 定性结果与缺陷定位

自动化飞行检查报告 (Automated Flight Inspection Report)

检查状态 (Status):
发现异常 (Anomaly Found): 256 个
A类异常数 (Class A Anomalies): 1 处

异常详情 (Anomaly Details):
异常 A1: 机身蒙皮雷击损伤 (Anomaly A1: Fuselage Skin Lightning Strike)
诊断信息 (Diagnostic Information):
严重等级 (Severity): A类 (Class A)
模型置信度 (Confidence): 98.5%
异常部件 (Component): 机身蒙皮 (Fuselage Skin)
具体位置 (Location): 机身左后侧, 45B段
三维模型坐标 (3D Coordinates): X=15.2, Y=3.5, Z=1.2
量化分析 (Quantitative Analysis):
初步分类 (Classification): 雷击烧蚀 (Lightning Strike Ablation)
损伤描述 (Description): 检测到一处由雷击造成的典型烧蚀点, 带有金属熔化痕迹, 损伤未穿透蒙皮。
几何测量 (Measurement): 损伤区域直径约 2.8 毫米, 最大深度 < 1 厘米。

处置与决策 (Disposition & Decision):
机长最终确认 (Captain's Final Confirmation):
同意并签派 (Confirm and Dispatch) 拒绝, 需进一步检查 (Reject, Requires Further Inspection)
提交决策 (Submit Decision)

Figure 5. User interface of the automated aerial vehicle inspection system

图 5. 自动化飞行器检查系统用户界面

除了定量的指标，定性分析也展示了我们系统的实用性。图 5 展示了系统用户界面的一个实例。在该案例中，模型成功检测到机身蒙皮上一个曾遭受雷击的微小痕迹，并精确地在交互式三维模型上进行了定位，同时给出了初步的损伤评估。

我们发现其他定性结果(未展示)也表明，我们的模型能够生成高精度的异常热力图，其高亮区域与地面真实缺陷掩码高度重合，无论是对于微小的裂纹，还是由液体泄漏引起的热异常，或是由凹痕引起的三维形变，我们的设计都实现了精确的定位。这些结果直观地证明了我们的模型在实际应用中的高精度定位能力和友好的用户交互设计。

3.6. 在公开数据集上的补充测试

为证明 MoE-ZSAD 算法核心思想的普适性，我们在两个相关的公开工业缺陷检测数据集上进行了补充测试：MVTEC AD 数据集(包含 15 类工业产品，如螺丝、瓶子等，总计约 5000 张图像)和 VisA 数据集(包含 12 类工业物体，如 PCB 板、管道等，总计约 10000 张图像)。这些数据集聚焦于表面异常检测，与 AeroSurf-AD 类似，但覆盖更广泛的工业场景。我们使用相同的模型架构，仅在正常样本上微调专家子模型(不使用任何异常样本)。结果显示，在 MVTEC AD 上的平均像素级 AUROC 为 0.94(优于 WinCLIP 基线的 0.89)，在 VisA 上的平均 AUROC 为 0.92(优于单一专家 ZSAD 模型的 0.87)。这些结果表明，MoE-ZSAD 的零样本学习和混合专家架构在异构工业对象上的泛化能力强，不仅限于航空领域，进一步验证了其在开放世界异常检测中的普适性。

4. 讨论

4.1. 方法对比分析

我们提出的 MoE-ZSAD 系统之所以取得优异性能，源于其关键技术组件的协同效应与深度整合。首先，DINOv3 视觉基础模型提供了一个强大而通用的特征空间。它通过在数十亿张无标签图像上的自监督学习，学会了理解图像的底层结构、纹理和高级语义，为后续的异常判断奠定了坚实基础。其次，零样本异常检测框架将问题从识别有限的已知缺陷转变为理解无限的正常模式，这是一种认知层面的范式转变。它从根本上解决了新发、未知缺陷的检测难题，使得系统具备了应对开放世界挑战的能力。最后，混合专家架构作为一种有效的归纳偏置，将“飞机是由不同视觉特性的部件组成”这一领域先验知识巧妙地融入模型设计中。这种分解式分析避免了模型试图学习一个模糊的、适用于所有部件的平均正常表征，而是为每个部件构建了精确、专门的正常模型，从而大幅提升了检测的灵敏度和准确性。

4.2. 细分结果分析

我们的消融研究结果深刻揭示了 MoE 架构的内在价值。在机身蒙皮这类视觉特征相对单一、平滑的部件上，单一专家模型与 MoE 模型的性能差距较小。然而，在起落架总成这类由大量不同材质、形状、光泽的子部件构成的复杂区域，单一模型难以建立一个统一的正常模型，导致性能下降。相比之下，MoE 架构中的起落架专家能够专注于学习该区域特有的复杂结构，从而实现了高达 7.7% 的性能提升。这表明，MoE 架构的优势在处理视觉异构性强的对象时尤为突出，这与飞机的结构特性高度契合。

4.3. 与现有研究的比较

我们发现，在自动化飞机检查领域，已有研究分别在机器人平台[6]或视觉检测算法[23]方面做出了贡献。然而，我们首次将高机动性机器人平台、全天候多模态感知以及专为航空领域复杂性定制的先进

MoE-ZSAD 架构进行端到端的系统性整合。这与依赖监督学习的传统自动化视觉检测方法相比，我们通过引入 ZSAD 范式，解决了传统技术无法检测未知缺陷的根本性瓶颈。同时，与通用的 ZSAD 方法(如 WinCLIP)相比，我们通过引入针对航空器异构特性的 MoE 架构，明显提升了模型在处理具有多种不同视觉特性部件的复杂对象时的性能[7]。因此，我们的研究通过提供一个经过验证的、能够解决关键未知异常问题的实用化系统，推动了该领域的技术发展。

4.4. 研究局限性

尽管我们在实验中展示了优异的性能和应用前景，但我们的研究仍存在一些局限性，这些局限性也为未来的研究指明了方向：

1. 数据方面的局限：模型的性能依赖于对正常状态的训练数据。对于一些可接受范围内的正常磨损、非标准的合规维修或涂装变化，如果未被充分包含在正常训练数据中，可能被误报为异常。
2. 方法/模型的局限：我们的 MoE 架构需要预先定义专家类别。对于一个全新的、部件划分完全不同的飞机型号，需要重新定义和训练专家集。
3. 资源限制：MoE-ZSAD 模型虽然在推理时是高效的，但其训练过程，特别是 DINOv3 基础模型的预训练，需要巨大的计算资源，这限制了模型的快速迭代和定制化。
4. 适用范围的局限：我们的系统专注于外部表面检查，无法检测飞机内部的结构疲劳、电子系统故障等问题。
5. 评估方面的不足：我们的实验是在构建的 AeroSurf-AD 数据集上进行的。尽管该数据集力求全面，但与真实机场运营中可能遇到的无限多样的光照、天气和遮挡组合相比，仍存在差距。
6. 动态环境适应性：我们的系统路径规划主要依赖基于静态三维模型的预规划路径，对于停机坪上临时出现的地面服务设备、人员等动态障碍物的实时、智能规避能力有待加强。
7. 内部缺陷检测：我们的研究采用的多模态传感器主要针对表面及次表面缺陷，对于材料内部的微裂纹扩展等深层结构健康问题检测能力有限。

4.5. 实际意义与应用价值

我们的技术方案将对航空运营产生深远且多维度的积极影响：

1. 理论贡献：我们在理论上验证了将大型视觉基础模型、零样本学习和混合专家架构相结合，是解决复杂工业场景中异构对象未知异常检测问题的有效途径，为相关领域的研究提供了新的思路。
2. 实践价值：
 - (1) 提升飞行安全：通过客观、量化的数据驱动决策，我们的系统能够发现人眼极易忽略的微小或新型缺陷，将安全隐患扼杀在萌芽状态，从而构建更坚实的安全防线[24]。
 - (2) 提高运营效率：我们将单次检查时间从人工的 10~15 分钟大幅缩短至 5~7 分钟，这能有效减少飞机在地面周转的时间，降低航班延误率，为航空公司带来直接的经济效益[25]。为增强该主张的可信度，我们提供了对单次检查时间的详细分解预算表(基于 A320 机型的模拟测试)：导航至观察点和路径跟踪(1~1.5 分钟)、机械臂定位与姿态调整(1 分钟)、多模态数据采集(1.5~2 分钟)、数据传输至服务器(0.5 分钟)、实时分析与异常检测(0.5~1 分钟)、报告生成与分发(0.5 分钟)。总耗时控制在 5~7 分钟内，该预算考虑了动态环境下的重规划延迟(不超过 0.5 分钟)，并通过高效的边缘计算优化实现。
 - (3) 推动数字化转型：我们的系统自动生成的、不可篡改的数字化检查报告为每一次飞行都创建了永久性的健康档案。这些海量数据可用于整个机队的健康状态监控、预测性维护计划的制定以及维修策略的优化，推动航空维修从传统的被动响应向现代的主动预测范式转变[26]。

4.6. 未来研究方向

基于上述局限性，我们在未来打算从以下几个方向展开研究：

1. 数据扩展方向：我们打算构建一个持续学习框架，使模型能够随着机队老化和维修历史的积累，动态更新其对正常状态的理解，以减少误报并适应变化。
2. 方法改进方向：我们将研究深度的多模态特征融合策略，在模型的特征层面进行可见光、红外、三维数据的早期或中期融合，而非仅作为并行的信息输入，从而进一步提升检测性能。同时，探索动态或自适应的 MoE 架构，使模型能够自动学习部件的划分。
3. 应用拓展方向：将我们的系统应用范围从起飞前检查拓展至更深入的定检维护环节，甚至集成更专业的无损检测(NDT)传感器(如超声波)，以检测飞行器内部结构损伤。
4. 验证与评估改进：在真实的、多变的机场运营环境中进行长期的、大规模的实地部署和验证，收集更多边缘案例，持续优化模型的鲁棒性。
5. 系统集成方向：将我们的系统与航空公司的维护、修理和大修(MRO)信息系统进行深度集成，实现从缺陷发现、工单生成到维修确认的全流程闭环管理。

5. 结论

我们的研究成功设计、实现并全面验证了一种用于飞行器起飞前检查的新型机器人化系统及方法。我们通过将自主移动机器人平台、全天候多模态传感器与创新的混合专家零样本异常检测(MoE-ZSAD)模型进行深度整合，系统性地解决了传统人工检查和现有自动化方案在主观性、覆盖范围、全天候作业能力以及对未知缺陷检测能力等方面的核心痛点。实验结果有力地证明，我们的系统在保证高效运行的同时，实现了对已知及未知表面缺陷的高精度检测和定位，其性能在各项关键指标上均超越了现有技术。这项工作不仅为提升飞行安全和航空运营效率提供了一个切实可行的技术方案，也为智能机器人在复杂工业检测领域的应用树立了新的标杆，同时代表了航空智能维护系统向着更加数据化、智能化、预测化的未来迈出了坚实的一步。

参考文献

- [1] Abbas, J. and Khare, A. (2024) Potential Errors during Final Inspection and Certification Process of an Aircraft Component. *Aircraft Engineering and Aerospace Technology*, **96**, 564-572. <https://doi.org/10.1108/aeat-11-2023-0294>
- [2] Sharov, V.D., Vorobyov, V.V. and Zatuchny, D.A. (2021) Probabilistic-Statistical Methods for Risk Assessment in Civil Aviation. Springer, 69-113. <https://doi.org/10.1007/978-981-16-0092-0>
- [3] Rajee Olaganathan, (2024) Human Factors in Aviation Maintenance: Understanding Errors, Management, and Technological Trends. *Global Journal of Engineering and Technology Advances*, **18**, 92-101. <https://doi.org/10.30574/gjeta.2024.18.2.0021>
- [4] Li, A., Qiu, C., Kloft, M., Smyth, P., Rudolph, M. and Mandt, S. (2023) Zero-Shot Anomaly Detection via Batch Normalization. arXiv: 2302.07849.
- [5] Sekelová, I., Korba, P., Koščáková, M. and Stanislavová, L. (2024) Examining Language Communication Difficulties in Aircraft Maintenance. *Acta Avionica Journal*, **26**, 63-72. <https://doi.org/10.35116/aa.2024.0008>
- [6] Piao, M., Wang, X., Wang, W., Xie, Y. and Lu, B. (2025) A Review of Robotic Aircraft Skin Inspection: From Data Acquisition to Defect Analysis. *Mathematics*, **13**, Article 3161. <https://doi.org/10.3390/math13193161>
- [7] Jeong, J., Zou, Y., Kim, T., Zhang, D., Ravichandran, A. and Dabeer, O. (2023) WinCLIP: Zero-/Few-Shot Anomaly Classification and Segmentation. 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Vancouver, 17-24 June 2023, 19606-19616. <https://doi.org/10.1109/cvpr52729.2023.01878>
- [8] Ren, J., Tang, T., Jia, H., Xu, Z., Fayek, H., Li, X., et al. (2025) Foundation Models for Anomaly Detection: Vision and Challenges. *AI Magazine*, **46**, e70045. <https://doi.org/10.1002/aaai.70045>
- [9] Koch, J., Jevremovic, D., Moenck, K. and Schüppstuhl, T. (2024) A Digital Assistance System Leveraging Vision

- Foundation Models & 3D Localization for Reproducible Defect Segmentation in Visual Inspection. *Procedia CIRP*, **130**, 387-397. <https://doi.org/10.1016/j.procir.2024.10.105>
- [10] Xu, J.C., Lo, S.Y., Safaei, B., Patel, V.M. and Dwivedi, I. (2025) Towards Zero-Shot Anomaly Detection and Reasoning with Multimodal Large Language Models. arXiv: 2502.07601.
- [11] Mu, S. and Lin, S. (2024) A Comprehensive Survey of Mixture-of-Experts: Algorithms, Theory, and Applications. arXiv: 2503.07137.
- [12] Vats, A., Raja, R., Jain, V. and Chadha, A. (2024) The Evolution of Mixture of Experts: A Survey from Basics to Breakthroughs. *Journal of IEEE Transactions on Artificial Intelligence*, **1**. <https://doi.org/10.20944/preprints202408.0583.v2>
- [13] Srivastava, P.A. (2025) Demystifying Sensor Fusion and Multi-Modal Perception in Robotics. *European Journal of Computer Science and Information Technology*, **13**, 76-90. <https://doi.org/10.37745/ejcsit.2013/vol13n267690>
- [14] Olufade, M.A., Bankole, E.A., Victor-Igun, O.O. and Junior, A. (2025) Multimodal Sensor Fusion for Autonomous Systems: Integrating Data from Various Sensors to Improve Decision-Making in Autonomous Vehicles and Robotics. *Journal of Basic and Applied Research International*, **31**, 37-54. <https://doi.org/10.56557/jobari/2025/v31i49405>
- [15] Costanzino, A., Ramirez, P.Z., Lisanti, G. and Di Stefano, L. (2024) Multimodal Industrial Anomaly Detection by Cross-Modal Feature Mapping. *2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Seattle, 16-22 June 2024, 17234-17243. <https://doi.org/10.1109/cvpr52733.2024.01631>
- [16] Tu, Y., Zhang, B., Liu, L., Li, Y., Zhang, J., Wang, Y., et al. (2024) Self-Supervised Feature Adaptation for 3D Industrial Anomaly Detection. In: Leonardis, A., Ricci, E., Roth, S., Russakovsky, O., Sattler, T. and Varol, G., Eds., *Computer Vision—ECCV 2024*, Springer, 75-91. https://doi.org/10.1007/978-3-031-72627-9_5
- [17] Huang, Z., Li, X., Liu, H., Xue, F., Wang, Y. and Zhou, Y. (2025) AnomalyNCD: Towards Novel Anomaly Class Discovery in Industrial Scenarios. *2025 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Nashville, 10-17 June 2025, 4755-4765. <https://doi.org/10.1109/cvpr52734.2025.00448>
- [18] Pan, Y., Wang, L., Chen, Y., Zhu, W., Peng, B. and Chi, M. (2025) PA-CLIP: Enhancing Zero-Shot Anomaly Detection through Pseudo-Anomaly Awareness. arXiv: 2503.01292.
- [19] Dou, S., Zhou, E., Liu, Y., Gao, S., Shen, W., Xiong, L., et al. (2024) LoRAMoE: Alleviating World Knowledge Forgetting in Large Language Models via Moe-Style Plugin. *Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, Bangkok, 11-16 August 2024, 1932-1945. <https://doi.org/10.18653/v1/2024.acl-long.106>
- [20] Lai, X., Yang, L., He, X., Pang, Y., Song, X. and Sun, W. (2023) Digital Twin-Based Structural Health Monitoring by Combining Measurement and Computational Data: An Aircraft Wing Example. *Journal of Manufacturing Systems*, **69**, 76-90. <https://doi.org/10.1016/j.jmsy.2023.06.006>
- [21] Liu, T., Blondel, M., Riquelme, C. and Puigcerver, J. (2024) Routers in Vision Mixture of Experts: An Empirical Study. arXiv: 2401.15969.
- [22] Jain, G., Hegde, N., Kusupati, A., Nagrani, A., Buch, S., Jain, P., Arnab, A. and Paul, S. (2024) Mixture of Nested Experts: Adaptive Processing of Visual Tokens. arXiv: 2407.19985.
- [23] Cao, Y., Zhang, J., Frittoli, L., Cheng, Y., Shen, W. and Boracchi, G. (2024) AdaCLIP: Adapting CLIP with Hybrid Learnable Prompts for Zero-Shot Anomaly Detection. In: Leonardis, A., Ricci, E., Roth, S., Russakovsky, O., Sattler, T. and Varol, G., Eds., *Computer Vision—ECCV 2024*, Springer, 55-72. https://doi.org/10.1007/978-3-031-72761-0_4
- [24] Ahmed, W. (2025) Artificial Intelligence in Aviation: A Review of Machine Learning and Deep Learning Applications for Enhanced Safety and Security. *Premier Journal of Artificial Intelligence*. <https://doi.org/10.70389/pjai.100013>
- [25] Moghadasian, S. (2025) AI-Powered Predictive Maintenance in Aviation Operations. *16th International Conference on Advanced Research in Science Engineering and Technology*, Chennai, 22-23 March 2026.
- [26] Sadeghi, A., Bellavista, P., Song, W. and Yazdani-Asrami, M. (2024) Digital Twins for Condition and Fleet Monitoring of Aircraft: Toward More-Intelligent Electrified Aviation Systems. *IEEE Access*, **12**, 99806-99832. <https://doi.org/10.1109/access.2024.3371902>