

北京邮电大学

项目报告（计划）书

项目名称：VIVIDIFY 低功耗智能超分相机

所在高校：北京邮电大学

所在学院：信息与通信工程

项目负责人：薛皓林

联系电话：13627601819

E-Mail：xhl518@bupt.edu.cn

项目成员：闫康、胡帅、罗玉杰

指导教师：景文鹏

2025 年 5 月 29 日

版权保护声明

- 1.本项目版权归项目负责人及其团队成员等共同所有；
- 2.未经项目负责人允许不得将本项目技术、模式等创意用作他用，特别是商业应用，否则将视为侵权。
- 3.借鉴或者引用本项目内容请注明来源及原作者。
- 4.对于不遵守此声明或者其他违法使用本文内容者，将依法保留追究权等。

一、项目背景

1.1 国家政策

我们的项目"VIVIDIFY 低功耗智能超分相机"深度契合国家战略规划，在技术创新和应用场景层面形成与政策导向的精准呼应。

绿色低碳：根据《2030年前碳达峰行动方案》的要求，电子设备需降低能耗并推广绿色设计标准，同时工信部《智能硬件产业创新发展专项行动》也明确提出鼓励智能摄像设备的低功耗设计。传统摄像设备因频繁捕捉和处理高分辨率图像，往往面临高能耗和散热问题，难以满足长时间稳定使用的需求。我们提出的双模摄像系统，采用“低功耗黑白视频流+间歇式高分辨率关键帧”架构，兼顾视频质量和系统节能。低功耗模式下，摄像头持续输出黑白低分辨率视频流，保证画面连贯性；高功耗模式则以低频率间歇性工作，捕获彩色关键帧，大幅减少数据处理量，有效降低了系统能耗。同时，黑白视频流的低功耗输出和云端服务器的超分重构处理，进一步优化了能耗表现。该方案创新性地实现了视频拍摄的高画质需求与低能耗长续航兼得，为电子设备的绿色设计提供了新的解决方案，完全符合国家绿色低碳政策对智能硬件的要求，有望成为未来智能摄像设备发展的新方向。

智能制造：《新一代人工智能发展规划》，明确将人工智能作为核心驱动力，提出构建自主可控的技术体系，重点支持智能感知、深度学习算法、边缘计算等方向。鼓励将人工智能技术与现有的关键技术相结合，实现人工智能的实际领域应用。本项目深度践行《新一代人工智能发展规划》"构建自主可控技术体系"的战略部署，我们采用的 NeuriCam 的超分重构模型，将先进的深度学习算法与实际应用领域深度结合，致

力于开发出满足市场需求且具有自主知识产权的创新技术。在智能感知方面，项目通过双模式摄像系统实现高效视频采集；在深度学习算法上，利用超分重构和注意力机制提升视频质量；在边缘计算领域，超分重构模型利用边缘设备的计算能力，实现了实时视频高效处理。我们实现了多种关键技术融合，不仅为在视频拍摄领域提供了新的人工智能技术解决方案，也为推动相关产业的智能化升级贡献了力量。

本项目具有显著的经济效益，在短视频创作领域，该系统能够为创作者降低约 60% 的设备成本（与主流 4K 摄像机相比），超过百万的内容生产者成为潜在客户，精准契合国家高质量发展的需求。随着短视频行业的爆发式增长，内容创作者对高性能、低成本拍摄设备的需求日益迫切。本项目通过创新的双模摄像系统，不仅大幅降低了设备成本，还提升了拍摄质量与创作灵活性，助力创作者产出更优质的内容，增强了市场竞争力。这不仅为创作者带来了实实在在的经济效益，也为整个行业注入了新的活力，推动了边缘计算，光学模组等产业的高质量发展，与国家推动产业升级和经济结构优化的战略方向高度一致。

1.2 技术/行业现状

短视频行业磅礴发展，短视频拍摄设备市场随之崛起：

2023 年，中国短视频行业在经历政策引导规范化、市场规模平稳化、内容生态成熟化、产业结构多元化的基础上，由爆发期、转型期进入了成熟期和深度发展期。在用户规模增长与产业链条完善的基础上，主流媒体短视频不断创新，网络微短剧等新兴形态兴起，短视频在垂直细分领域逐渐深入并与文旅产业有效结合，推动城市形象广泛传播。在政策、用户与技术的多重影响下，短视频行业的商业化进程也在不断加速。

短视频用户规模持续扩大，成为推动行业增长的重要因素。中国互联网络信息中心发布的第 54 次《中国互联网络发展状况统计报告》显示，截至 2024 年 6 月，我国网络视频用户规模已接近 11 亿大关，达到 10.9967 亿人，其中短视频用户规模达 10.68 亿人，较 2022 年同期增长 6761 万人，占网民整体的 97.1%，从下图中，我们可以清晰地看到中国互联网网民规模和互联网普及率的历年变化，以及网络视频用户的增长和占比情况。这些数据揭示了近年来短视频市场的迅猛扩张。



图 1.2-1 全国网民规模和互联网普及率



图 1.2-2 全国短视频用户规模

短视频行业自 21 世纪初兴起，至 2025 年，全球市场规模已达到百亿美元级别。短视频平台为网络用户提供了丰富的内容供给、多元的场景体验，从而推动网络用户规模的持续增长，并带动整个行业规模的扩张。据 Mob 研究院的调查数据，2023 年短视频市场规模近 3000 亿元，

在涵盖内容创作、网络音频、网络直播、综合视频等的网络视听行业市场，短视频市场占比达到 40.3%，市场规模为 2928.3 亿元。

国内市场爆发：根据相关调研数据，预计到 2050 年，短视频行业的市场规模将实现数倍增长。以我国为例，根据最新发布的《2025 年中国短视频行业研究报告》显示，2025 年我国短视频**市场规模将达到 500 亿元美元**，占全球市场份额的 20%。在未来 25 年内，这一比例有望进一步提升。

创作者激增：随着短视频行业的爆炸性增长，短视频内容创作者也如雨后春笋般涌现。以抖音、哔哩哔哩为代表的平台，**2020 年，仅抖音一个平台，创作者总量已经突破惊人的 1.3 亿**，且保持着年均 12% 以上的高速增长。在竞争白热化的创作生态中，内容创作需求正呈现井喷式爆发。

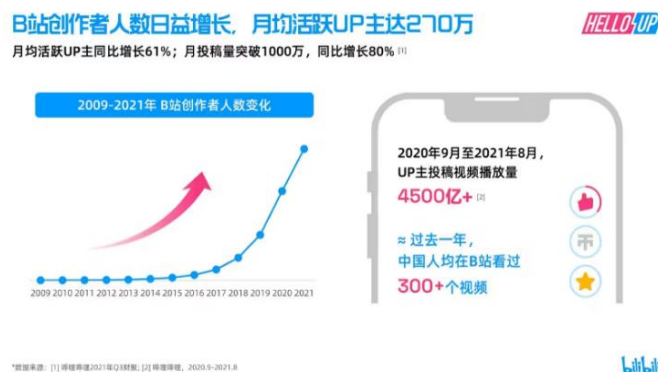


图 1.2-3 21 年哔哩哔哩创作者总量

全民创作的趋势在抖音生根发芽，每个人都充满了无限可能性

全民创作在抖音



图 1.2-4 20 年抖音创作者总量

市场竞争加剧：虽然用户规模持续增长，但其增速逐步放缓，用户增长的红利逐渐消退，行业真正步入存量竞争的时代。在用户规模增长放缓的同时，短视频用户黏性不断增强，2023 年短视频人均单日使用时长达到 151 分钟，短视频应用平台的头部效应尤为凸显，抖音、快手两大短视频平台占整体份额的 95.3%。随着用户对网络视听内容产品的需求变化，高质量、精品化、个性化的内容生产成为短视频发展的未来方向。短视频为网络视听行业的转型升级提供了新的路径，成为媒体深度融合的重要力量。短视频的精品化转向是必然趋势。短视频内容新颖，高质量，精品化才是提高竞争力的关键。

内容创作领域正在经历一场前所未有的变革。我们正处于一个最佳的时期，也是最具挑战的时期。尽管内容消费市场持续扩大，但内容创作的竞争却变得愈发激烈。粉丝超过 1000 万的博主在主流平台上的占比仅为 0.01%，这一数据反映了内容创作领域的马太效应。

短视频行业的爆发式增长，短视频创作者数量急剧增加，短视频行业竞争加剧，短视频内容高质量要求，持续推动拍摄设备需求升级。

2025 年全球短视频拍摄设备市场规模预计突破 480 亿美元，中国贡献超 35% 份额，2025 年中国摄影器材行业市场规模持续扩大，展现出强劲的

增长动力。随着国民经济的快速发展和消费升级趋势的深化，创作者对高品质摄影设备的需求日益增长。硬件创新与场景细化成为行业核心驱动力。设备轻量化、智能化趋势显著，4K/8K 超高清、HDR 成像、多镜头协同等技术渗透率较 2023 年提升 22 个百分点，用户对便携性与专业功能的双重需求倒逼产业迭代。

因为短视频创作者数量激增和短视频创作的竞争压力，创作者对摄影器材的需求从基本满足转向追求高品质和体验。高端摄影器材市场逐渐扩大，创作者愿意为更好的成像效果和拍摄体验支付更高的价格。这一趋势推动了摄影器材行业的产业升级和技术创新，促进了高端产品的研发和上市。

传统相机虽在图像质量、光学变焦等方面表现优异，但在动态场景捕捉方面存在明显局限。面对户外运动、夜景等复杂场景，传统相机常因高速运动导致画面模糊或低光环境下噪点过多而难以获得理想的拍摄效果。尽管部分高端机型通过提升感光度、优化防抖功能来改善，但整体提升有限，未能彻底解决痛点。而短视频拍摄场景往往都是随时随地的不固定的动态场景，特别是旅行博主和 vlog 博主的使用场景，博主的拍摄场景在不断的切换，要想拍出可用的高质量素材，对拍摄设备在动态场景捕捉要求更高。

随着短视频行业的兴起，用户对于高性价比、易用性强且具备专业级视频拍摄功能的设备需求日益增长，传统相机复杂的操作流程和较高的价格门槛使其难以满足这一庞大用户群体的需求。

本项目推出的超分相机，精准定位用户需求，全面支持高分辨率（4K/8K）视频拍摄，并凭借卓越的动态场景捕捉能力，无论是户外极限运动的高速场景，还是夜间低光环境下的细腻画面，均能轻松驾驭，

为创作者提供专业级的视频质量。这不仅满足了当下市场对高品质视频内容的迫切需求，更为创作者在激烈的竞争中脱颖而出提供了有力的技术支撑，市场潜力不可限量。有着巨大的发展潜力。

1.3 技术/行业痛点

目前短视频拍摄设备主要面临如下几个方面的问题：

1.3.1 续航问题：

一些高性能的拍摄设备，如专业摄像机、无人机等，为了保证拍摄高质量和功能的多样性，往往需要有着较高的能耗，续航时间较短，无法满足长时间拍摄的需求。4K/8K 视频录制、AI 实时处理（如夜景优化、人像追踪）大幅增加功耗。手机拍摄 10 分钟即过热降频，运动相机持续录制时间仅在 1 小时之内，而在一些户外或偏远地区拍摄时，充电条件有限，设备的电量不足问题会进一步影响拍摄进度和创作体验，进一步限制了设备的使用，给创作者带来不便。



1.3.2 动态场景捕捉能力弱：

对于高速运动的物体或场景，如户外运动等高速移动的场景，传统拍摄设备在捕捉这类快速变化的画面时经常会出现拖影、模糊等现象，无法精准地记录下每一个瞬间，难以满足用户对于动态场景高质量拍摄的要求。这种局限性限制了创作者在拍摄精彩动态画面时的发挥，也影响了最终视频素材的质量和专业性。



图 1.3-1 高速移动场景

1.3.3 画质与便携性的矛盾：

短视频创作者通常更倾向于使用便携性高、操作简便且功能全面的视频拍摄设备，以便在各种复杂场景中迅速捕捉到高质量的动态画面，同时降低设备成本和操作难度，但是如手机、运动相机等小巧便携的拍照设备，受限于小尺寸传感器，在暗光场景下噪点控制、动态范围表现不足，难以满足高质感拍摄需求。然而，传统的相机虽搭载大传感器和外接镜头，但是却由此大多体积庞大、**操作复杂**，牺牲了移动灵活性，并且多数机型的设计主要侧重于拍照功能，视频功能相对较为薄弱，难以满足短视频创作者对于高效、灵活视频创作的需求。



图 1.3-2 体积大行动不便

1.3.4 设备成本高：

高质量的短视频拍摄设备，像专业级摄像机、高端镜头、稳定器、无人机等，价格昂贵，对于众多创作者，尤其是个人创作者和小型团队而言，是一笔相当大的开支。此外，随着科技的不断进步和更新换代，拍摄设备的更新周期也在缩短，这无疑进一步加重了创作者的经济负担，限制了一些创作者在设备投入上的选择和升级。

二、项目介绍

2.1 产品概述

我们的产品是一个与基于关键帧的超分重构模型相结合的创新性的双模式 4K 摄像机，VIVIDIFY 低功耗超分相机

2.1.1 产品来源概述

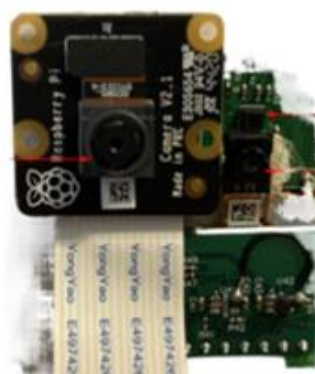
随着物联网技术的迅猛发展，短视频已经成为我们日常生活中不可或缺的一部分。对于短视频博主，尤其是旅行博主，vlog 博主，户外博主来说，他们需要长时间拍摄高分辨率、高帧率的视频素材，这不仅能提升内容的吸引力，也是创作高质量作品的关键。然而，这样的拍摄需求往往伴随着高功耗的问题，导致电池续航时间缩短。在户外复杂不断变换的场景中，电源并不总是容易获得，这限制了博主们的创作自由度。现有摄像头技术在满足高质量视频捕获的同时，往往难以实现低功耗运行。这意味着在追求高分辨率和色彩表现的同时，博主们不得不面对电池续航的挑战。受限于硬件和成本，现有解决方案在提供长时间续航和高质量视频输出之间难以找到平衡点。

所以，我们希望研发一款创新性的产品，基于关键帧的视频超分双模式摄像系统，采用“低功耗黑白视频流+间歇式高分辨率关键帧”架

构，兼顾视频质量和系统节能。该产品专为短视频创作者设计，特别是那些经常在户外拍摄的旅行、生活以及其他领域的博主。系统利用双模式摄像的模式结合神经网络超分重构技术，不仅确保了视频内容的超高清分辨率，而且因为双模系统带来的低功耗，还大幅延长了拍摄时长，让博主们在创作过程中能够维持高标准的艺术表现力，并享受到更加自由灵活的拍摄体验。这种灵活性不仅扩展了创作的边界，还为探索新的创作可能性提供了广阔的空间。最终形成了我们的"VIVIDIFY 低功耗智能超分相机"。



预计成品图



内置双摄像头
无线传输模块

图 2.1-1 VIVIDIFY 低功耗智能超分相机成品图

2.1.2 产品创意来源

我们的项目创意主要源自以下几个方面的深入思考：

首先，我们想到视频流中的关键帧通常包含丰富的细节信息，这些信息对于提升整体视频质量至关重要。因此，我们考虑通过算法处理，利用视频超分技术，将这些高分辨率的关键帧与低分辨率的连续视频帧相结合，实现视频的超分辨率和色彩化。

其次，现有的深度学习方法在视频超分辨率方面已取得了显著成果。但这些方法在低功耗物联网摄像场景中的应用，往往面临计算量大、功耗高的难题。我们在借鉴这些方法优点的基础上，进行了针对性的改进和优化，以适应低功耗场景的需求。

最后，硬件设计对于实现低功耗和高性能同样至关重要。我们通过优化硬件结构和算法实现，使得整个系统在保证视频质量的同时，能够显著降低功耗，找到提供长时间续航和高质量视频输出之间的平衡点，以满足长时间拍摄的需求。

2.1.3 产品架构

下图是我们的产品整体架构示意图：



图 2.1-2 整体架构

应用层 – 边缘计算云服务器实现超分重构

在应用层，系统专注于实现高分辨率彩色视频的重建。这一层次集成了超分重构模型，并将其部署在边缘计算云服务器中，能够接收低分辨率黑白视频流和以固定频率捕获的高分辨率彩色关键帧，通过先进的

基于关键帧超分重构模型，输出高质量的 4K 彩色视频。这一层次是系统对外呈现的最终产品，直接面向最终用户的需求。

软件层 - 神经网络模型与处理单元

神经网络模型是系统的核心，负责整体的视频处理流程，包括超分辨率和彩色化。包含多个关键组件：特征提取器，从视频帧中提取关键特征，为后续处理提供基础。双向循环传播网，利用视频帧之间的时间关联，增强模型对动态场景的理解。光流对齐和变形卷积，来处理视频中的高速运动情况，确保动态条件下视频输出的高清晰度和准确性。注意力特征过滤：根据特征的相关性动态调整权重，优化视频质量。

该层确保视频数据在经过一系列复杂的处理步骤后，能够呈现高分辨率和准确的色彩的高质量视频。

数据层 - 多源视频数据集成

数据层负责管理来自不同摄像机的数据流包括：低分辨率、黑白视频流，保障超分后视频的连续性。高分辨率、彩色关键帧，提供关键帧以辅助色彩和细节的重建。

数据层需要高效地处理和同步这些数据流，并通过 TCP 连接将处理后的数据传输到软件层，为软件层提供准确的输入。

硬件层 - 摄像机系统与功耗管理

硬件层包括两个子系统：

低功耗摄像系统：低功耗摄像头持续输出 30fps 黑白低分辨率视频流，确保视频的连续性和流畅性；**高功耗摄像系统：**高功耗摄像头间歇捕获 1fps 高分辨率彩色关键帧，提供还原细节。硬件层的设计注重功耗

管理，通过智能调度和占空比控制，实现能效最优化。这种设计允许系统在保持低功耗的同时，通过关键帧的补充，实现高质量的视频输出。

我们的摄像系统主要分为两部分：双模式摄像系统，由两个协同工作的摄像头（“低功耗黑白视频流+间歇式高分辨率关键帧”）和边缘计算云服务器（部署着我们的超分重构模型）组成。在拍摄时两种模式的摄像头同时运行。

低功耗模式下摄像头持续运行，连续输出的每秒 15 帧黑白低分辨率视频流，保证了画面的连贯性。

而高功耗模式下的摄像头则以每秒一次的频率间歇性工作，捕获高分辨率的彩色关键帧。在拍摄过程中，系统会将拍摄的黑白视频流和关键帧的图像数据通过 **TCP 连接**，传输到边缘计算云端服务器进行实时超分重构。通过基于关键帧的超分重构模型，利用高分辨率的彩色关键帧对黑白低分辨率视频流进行超分操作，实时输出高分辨率的视频。

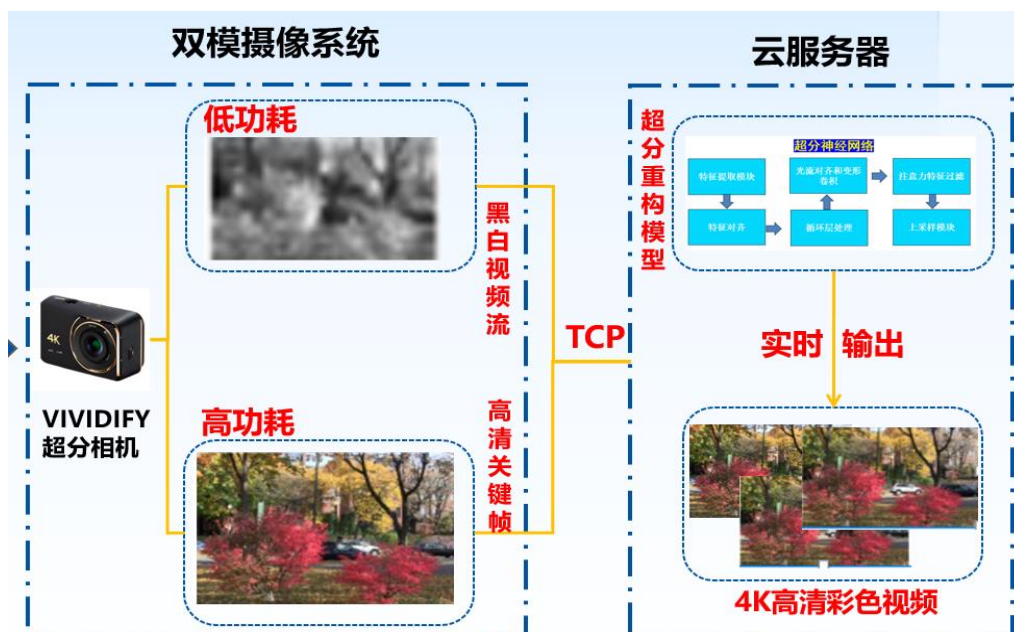


图 2.1-3 运行流程示意图

2.2 产品技术

双模式摄像系统：我们创新性地设计了一种双模式摄像系统，该系统融合了低功耗和高功耗两种工作模式。低功耗模式的摄像头以其持续性工作提供连续的黑白、低分辨率视频流，而高功耗模式的摄像头则以间歇性工作的方式，每秒捕获一次高分辨率和彩色的关键帧，实现了对重要低分辨率情况下图像缺失的高频细节精准捕捉，以及图像色彩的获取。

实时神经网络解码：我们的系统在网关设备上实现了实时神经网络解码。该接入点或路由器接收来自双模式摄像系统的黑白低分辨率视频流，同时结合高分辨率彩色关键帧，通过实时解码器重建出高分辨率彩色视频。这一过程不仅保持了系统的低功耗特性，同时也确保了视频输出的高质量。

注意力特征滤波器机制：我们引入了一种新颖的注意力机制，该注意力特征滤波器机制是一项突破性创新，它通过自适应权重分配和跨帧关联性学习，显著提升了视频处理的智能化和精确度。该机制不仅实现了灰度到彩色的精确映射，增强了细节特征的提取，还有效抑制了噪声，提升了视频的色彩丰富度和真实感。它通过上下文感知和运动补偿，优化了对快速运动物体的处理，并通过多尺度特征融合，在不同分辨率层次上捕捉细节，进一步提升了视频质量。同时，该机制在设计上平衡了计算效率与性能，具备良好的泛化能力，适用于多样化的场景和条件。

双向循环网络架构：双向循环网络架构允许模型在视频重建过程中同时考虑过去和未来的帧，从而在时间维度上实现了全面的信息利用。通过深度整合前后帧的上下文信息，显著提升了对场景的理解。这种深

度整合不仅包括视觉信息，还包括运动信息和可能的语义信息，为视频重建提供了丰富的数据支持，使得视频的每一帧都能够更在更宽广的上下文中被理解和处理。对于动态场景，尤其是快速运动的物体，双向循环网络能够更好地捕捉和预测物体的运动轨迹，从而在超分辨率重建时保持物体的连贯性和清晰度。

光流对齐和变形卷积：使用光流对齐和变形卷积技术来处理视频中的高速运动情况，确保动态条件下视频输出的高清晰度和准确性。光流估计技术精确追踪物体运动轨迹，与变形卷积的自适应卷积核相结合，实现快速移动物体的特征对齐和细节增强，减少运动模糊。这些技术优化了运动补偿，提升了系统在复杂场景下的鲁棒性，并保障了实时处理性能。它们适用于多种场景，与深度学习模型协同，增强了系统的智能化水平。

固定单应性变换：有效解决了多摄像头视角差异问题。这项技术通过几何校正和视角融合，确保了不同摄像头捕获的图像在数据融合时的一致性，减少了视觉误差，增强了场景理解。它优化了视觉呈现，提高了自动化处理效率，适用于室内外、不同光照和动态场景。固定单应性变换还提升了用户体验，尤其在多摄像头直播或虚拟现实，提供了沉浸式观看体验。此外，它与系统中的其他先进技术如光流对齐、变形卷积和注意力机制等集成，共同提升了视频处理的整体性能，展示了视频处理领域中数学和算法创新的实际应用，为多摄像头视频处理提供了高效、可靠的解决方案。

2.3 产品优势

2.3.1 解决“高清画质与续航不可兼得”的功耗痛点

传统拍摄设备面临着高功耗困境，传统相机拍摄高分辨率视频需自身持续处理大量数据，导致功耗激增。创作者想要拍摄出高质量的视频素材就必须舍弃长时间拍摄的可能。特别户外拍摄时需要频繁更换电池或中断创作。据实测，4K 拍摄模式下手机续航普遍低于半小时，甚至有的安卓手机只能拍摄十几分钟视频，变回自行退出拍摄，4K 相机续航低于一个小时。现有低功耗方案（如 AOV 技术）为省电采用“1 秒 1 帧”的极低帧率，虽延长续航但牺牲流畅度，高速运动场景漏拍率达 30% 以上。对于想要持续拍摄高质量，高分辨率，流畅的创作者而言，这些都是致命的缺点。

“双模协同”解决续航痛点，我们提出 VIVIDIFY 的创新方案，利用双模式协同设计，低功耗摄像头持续输出 30fps 黑白低分辨率视频流，保障基础流畅性；高功耗摄像头间歇捕获 1fps 高分辨率彩色关键帧，提供还原细节，双模式设计显著降低设备功耗。同时我们在边缘计算云服务器上部署了我们的超分重构模型。VIVIDIFY 摄像机本身并不需要自己进行视频处理，只需要将视频数据传输到边缘计算服务器，在边缘计算服务器中，会进行基于关键帧的视频超分流程，关键帧与黑白视频流在云端实时融合生成高清 4K 视频，进一步降低自身功耗。根据我们的实际测试，VIVIDIFY 低功耗摄像机设备端功耗降低了 70% 以上，续航延长 5-7 倍，可以达到 7 个小时的高强度的长时间拍摄，并在云端实时输出拍摄的 4K 视频。

因此，VIVIDIFY 低功耗摄像机完美的解决了“**高清画质与续航不可兼得**”的功耗痛点，为视频创作者提供了更好的创作环境，也提供了无限的创作可能。

2.3.2 解决“动态场景模糊与细节失真”的画质痛点

传统摄像行业存在一个通病：算法局限导致伪影与失，传统相机和手机在高倍变焦时易出现文字扭曲、人脸失真等问题，因此依赖于 AI 超分技术生成式补全，非真实光学数据，这个过程不但会增加设备的功耗，同时也会造成图像失真。同时传统相机行业还面临一大问题，运动模糊难以消除，拍摄高速运动的物体会导致重影，非常影响画面质量，以及素材质量。存在着“**动态场景模糊与细节失真**”的画质痛点。

“超分重构”解决画质痛点，VIVIDIFY 低功耗超分像机的在这方面取得突破性技术，多算法融合架构，我们在基于关键帧的超分模型中，融入了注意力特征滤波器机制，动态分配关键帧权重，精准还原色彩与高频细节，减少画面失真；我们利用光流对齐+变形卷积，彻底解决高速运动模糊，位移误差控制在 $<5\%$ ，极大程度上的减少了画面模糊，杜绝了重影现象；同时我们构建了双向循环网络，超分时融合前后帧上下文信息，进一步提升运动连贯性。

工业级精度：在户外奔跑测试中，毛发纹理、文字标识等细节还原度达 98%。彻底解决了“**动态场景模糊与细节失真**”的画质这一痛点

2.3.3 突破“应用复杂，携带不便”的痛点

传统相机通常功能繁多，设置复杂，需要用户具备一定的摄影知识和经验才能熟练操作。对于非专业用户来说，学习如何正确设置拍摄参数、调整镜头光圈、快门速度等非常耗时，增加了使用难度。而且传统相机体积较大，特别是专业级单反相机，镜头和机身重量重，携带起来

非常不便。这给用户在移动拍摄过程中增加了负担，尤其在需要频繁移动或在复杂地形中拍摄时，如户外拍摄或旅游 vlog，用户不得不携带多个镜头和配件，进一步增加了使用上的不便，存在“应用复杂，携带不便”的痛点



图 2.3-1 佳能专业级单反摄像机的体积

“小巧轻便”解决携带痛点，VIVIDIFY 低功耗超分摄像机配备了直观且用户友好的界面，用户无需具备复杂的专业知识，即可快速上手进行拍摄，降低了使用门槛。VIVIDIFY 低功耗超分摄像机能拍摄与专业单反摄像机画质一样的视频，不需要特别专业的镜头和其他摄像装备，所以 VIVIDIFY 超分相机机身小巧轻便，易于携带，易于操作，特别适合在移动拍摄过程中使用。用户可以轻松地将其放入背包或口袋中，随时随地进行拍摄。在面临各种复杂地形和环境中拍摄时更加灵活，无需担心设备的重量和体积限制。突破了“应用复杂，携带不便”的痛点。

2.3.4 解决“高清画质与低价格不可兼得”的价格痛点

传统高清摄像机，尤其支持 4K/8K 分辨率的专业设备，因采用高品质传感器、镜头及图像处理器，成本高，一般价格在 8000 左右，价格动辄数万元，对个人创作者和小型团队都是沉重经济负担。低价摄像机

为控制成本，常在关键部件上妥协，如采用小尺寸传感器、低性能镜头和简陋图像处理器，致使其在分辨率、色彩还原、低光表现等关键画质指标上表现欠佳，无法满足高质量视频创作的需求。存在“高清画质与低价格不可兼得”的价格痛点。



图 2.3-2 索尼 4K 摄像机价格

“超高性价比”解决价格痛点，VIVIDIFY 摄像机采用创新的双模摄像系统和超分重建技术，在保证高清画质的同时，大幅降低设备成本。VIVIDIFY 摄像机并不需要专业 4K 单反相机一样的高品质传感器、镜头及图像处理器（本地不需要处理图像数据），所以 VIVIDIFY 摄像机能大幅降低成本，其价格显著低于传统专业级 4K/8K 摄像机，使更多用户能够以较低的价格获得高质量的拍摄设备。通过“低功耗黑白视频流+间歇式高分辨率关键帧”架构和先进的超分重建算法，VIVIDIFY 摄像机能够输出高质量的 4K/8K 视频。其动态场景捕捉能力和低光性能出色，确保在各种复杂环境下均能提供清晰、细腻的画质。

2.3.5 突破“应用场景单一”的兼容性痛点

传统相机在应用场景上存在明显局限性。消费级相机因依赖持续供电，难以在野外无电环境长期部署，限制了其在农田监控、野外科考等场景的应用。工业相机虽具高精度，但高功耗与高昂成本（如 3D 视觉

设备超 2000 美元) 使其仅限于特定专业领域, 难以普及推广, 应用场景受限。存在“应用场景单一”的兼容性痛点

“应用场景多样”解决兼容性痛点, VIVIDIFY 摄像机凭借功耗革新, 实现多场景适配。其低功耗设计使其太阳能供电即可支持 7×24 小时运行, 完美适配无电环境下的农田监控和野外相机等场景。在画质保障方面, 超分重构技术确保视频细节清晰, 满足安防领域的人脸、车牌识别以及工业质检中小至 0.1mm 缺陷检测的需求。同时, 成本优势显著, 硬件成本较传统设备降低 70%, 通过算法优化取代部分高规格光学元件, 兼顾性能与经济性。



图 2.3-3 智能农业

野外相机

2.3.6 总结：技术革新与行业价值

VIVIDIFY 通过“双模式硬件架构+云端超分算法+多传感器融合”三位一体设计, 实现了五大革命性突破:

功耗革命: 重新定义高清拍摄的能效边界, 续航提升 5-7 倍;

画质跃迁: 攻克动态模糊与 AI 失真, 细节还原度>98%;

系统鲁棒性: 视角自动校正, 复杂环境稳定性达工业级标准;

便捷性：体积大大减小，自身重量减小，操作简单适应复杂环境；

场景普适性：从短视频创作到物联网监测，成本下降 70%

2.4 关键技术

2.4.1 双模式摄像系统

双模式摄像系统由两个协同工作的摄像头和一个作为数据处理核心的接入点或云服务器组成。系统巧妙地结合了低功耗和高功耗两种运行模式，内置两个摄像头分别执行不同模式，在拍摄时两种模式同时运行。工作在低功耗模式下摄像头持续运行，以每秒 15 帧的帧率输出连续的黑白低分辨率视频流，保证了画面的连贯性。而高功耗模式下的摄像头则以每秒一次的频率间歇性工作，捕获高分辨率的彩色关键帧。在拍摄过程中，系统会将拍摄的黑白视频流和关键帧传输到云端服务器进行超分重构，实时输出高分辨率的视频，这个过程不会增加系统本身的功耗。

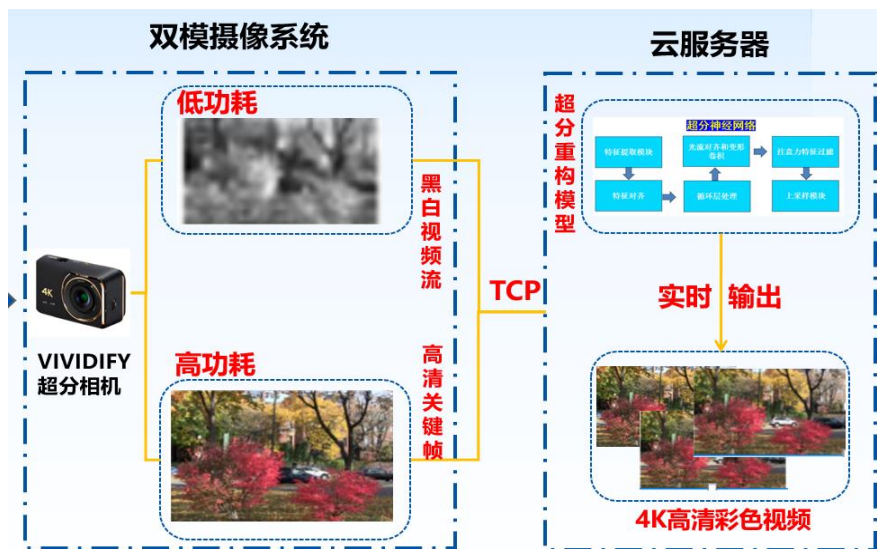


图 2.4-1 双模摄像系统工作架构

与传统 4K 相机比，为获得高画质，相机需要不断地捕捉和处理高分辨率的图像内容。需要处理大量的数据，因而导致相机的功耗显著增

加，相机续航时间减少，无法实现长时间的持续使用。而双模式摄像机以更低的频率获取高分辨率图像，大大减少了摄像机的数据处理量，而拍摄低分辨率视频功耗也很低，超分重构过程也不会额外增加相机本身的功耗。在这样三重保险下，双模式摄像系统对比于普通相机，在拍摄同样的高分辨率视频时，相机本身的功耗会大幅降低并且大幅度延长了相机的工作时间。双模式摄像系统的设计不仅精准地捕捉了在低分辨率视频中缺失的高频细节，还实现了图像色彩的准确获取，从而在保持能源效率的同时，显著提升了视频的质量和细节表现。

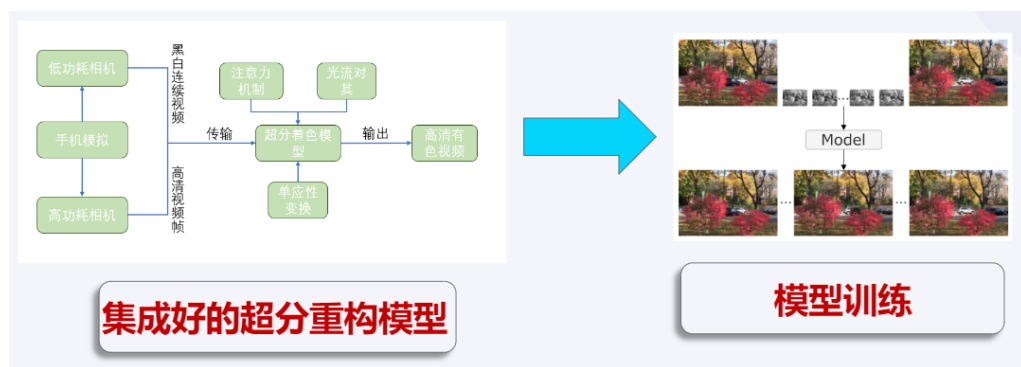


图 2.4-2 双模摄像效果实现

从上图可见，系统通过双模式摄像机同时采集低分辨率连续帧（黑白图像）与间歇性的高分辨率关键帧（彩色图像），输入至所设计的神经网络模型中。模型通过特征提取、时序融合与超分重建等多个子模块，对输入序列进行联合建模与像素级增强。最终，系统能够将原始低分辨率灰度视频流与稀疏的彩色关键帧重构为帧帧连续、细节丰富、色彩真实的高分辨率视频序列。该结果充分验证了双模视频采集策略与深度学习超分模型的有效结合，为低功耗高质量视频重建提供了可行路径。

2.4.2 超分神经网络

我们的系统在网关设备上部署了高效的实时超分重构。系统能够接收来自双模式摄像系统的黑白低分辨率视频流，并与高分辨率彩色关键帧相结合，通过超分重构实时重建出高分辨率的彩色视频。这一创新流程不仅维持了系统的低功耗特性，而且还确保了视频输出的卓越画质，为用户提供了既节能又高清晰度的视频拍摄体验。

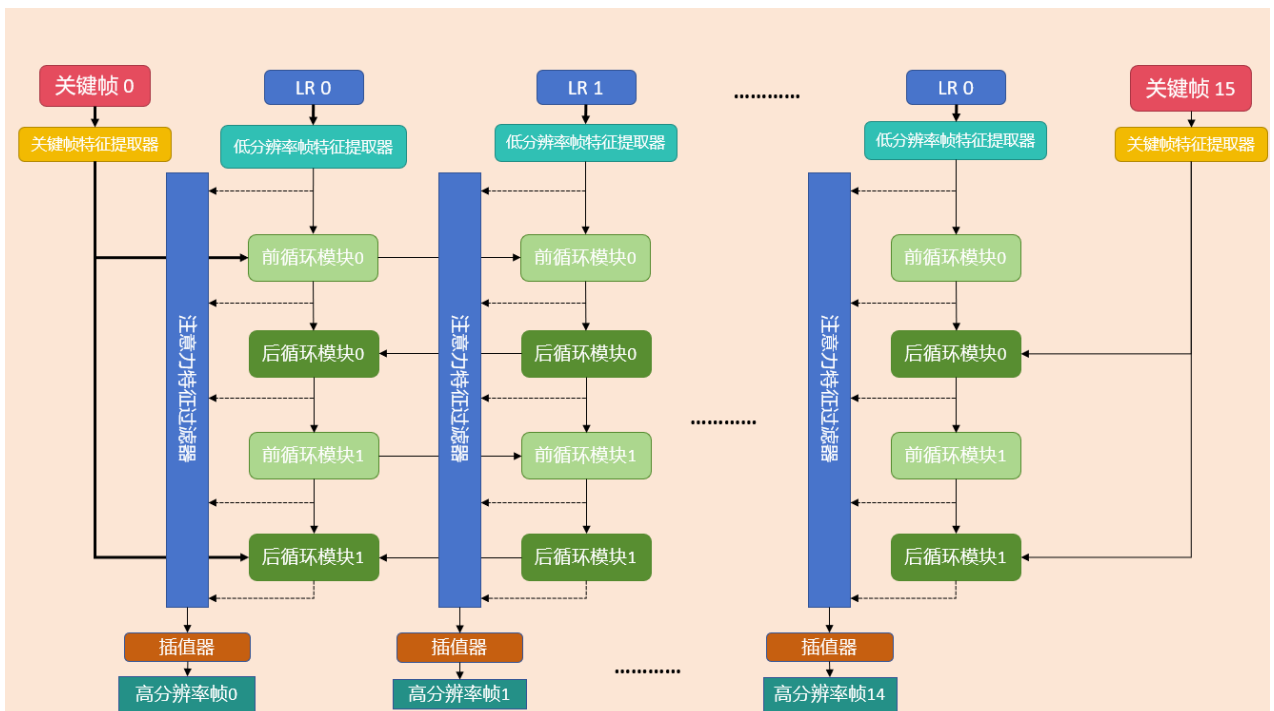


图 2.4-3 超分神经网络系统总架构

如上图所示，为了更好的还原图像细节，在超分辨率重建时保持视频的连贯性和清晰度。我们构建了双向循环网络架构，这个架构能够允许我们的模型在视频重建过程中同时考虑过去和未来的帧，，从而在时间维度上实现了全面的信息利用。双向循环网络架构通过深度整合前后帧的上下文信息，显著提升了对场景的理解。这种深度整合不仅包括视觉信息，还包括运动信息，为视频重建提供了丰富的数据支持，使得视频的每一帧都能够在更宽广的上下文中被理解和处理。对于动态场景，尤其是快速运动的物体，双向循环网络能够更好地捕捉和预测物体的运

动轨迹，从而在超分辨率重建时保持物体的连贯性和清晰度，实现效果见下图。

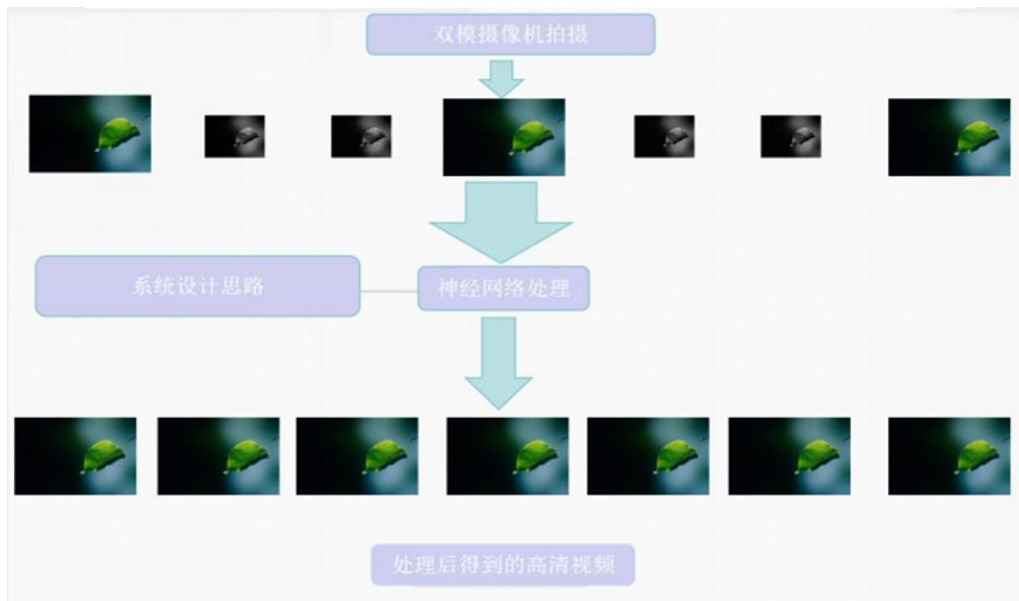


图 2.4-4 超分神经网络系统输出效果图

接下来，将从网络结构的各个关键模块出发，详细介绍本项目针对高效视频超分辨率任务所设计的核心组件，包括轻量化残差特征提取模块、光流估计模块及二阶可变形对齐机制、特征注意力聚合模块。各模块分别面向图像空间特征表达、时序特征对齐与复杂运动场景下的空间预测展开，协同提升了整体系统的重建性能与泛化能力。下面将对各模块的设计思路、具体结构及其性能优势进行详细阐述：

(1) 双路残差特征提取模块设计

针对关键帧与黑白视频帧的差异性，我们引入双路残差特征提取模块。关键帧通过下采样与深层残差卷积模块提关键细节特征，而普通帧路径以残差学习为基础，提取局部结构与纹理信息。通过对两类帧采用差异化处理策略，该模块不仅增强了整体特征表达能力，也在保持性能的同时显著降低了模型参数规模与计算开销，为后续的时序传播与重建过程奠定了高效且稳定的特征基础。

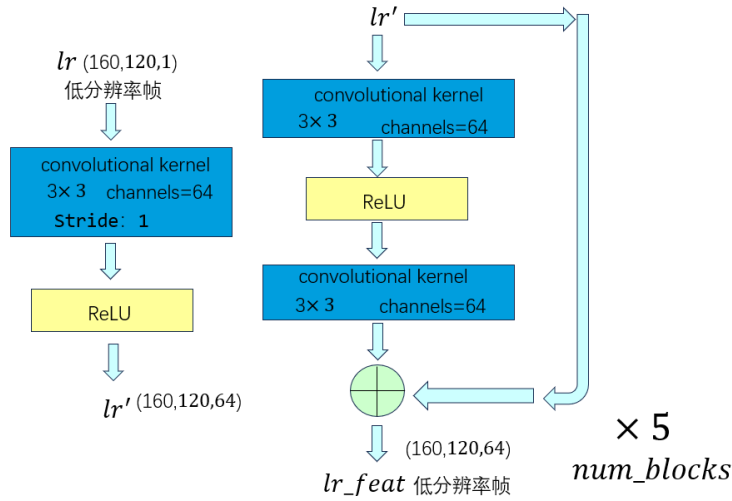


图 2.4-5 低分辨率帧特征提取模块具体设计

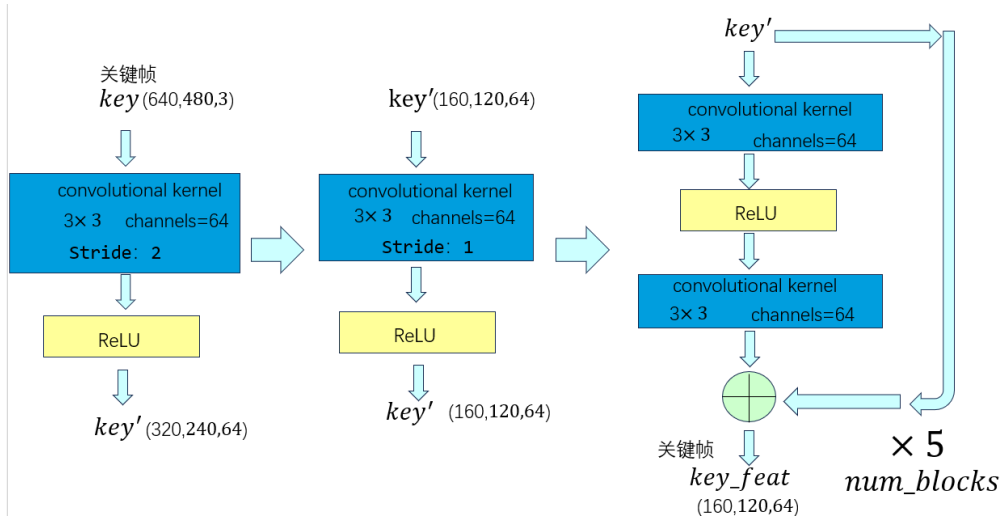


图 2.4-6 关键帧特征提取模块具体设计

在具体结构上，关键帧输入图像首先经过一个步长为 2 的 3×3 卷积操作，输出通道数为 64，完成初步下采样，输出尺寸为 $320 \times 240 \times 64$ 。随后，特征图经由 ReLU 激活函数引入非线性表达能力，再通过另一个步长为 1 的 3×3 卷积与 ReLU 激活进一步提取语义特征，最终得到尺寸为 $160 \times 120 \times 64$ 的关键帧特征表示。在此基础上，我们堆叠了 5 个残差卷积块（Residual Blocks），每个块包含两个 3×3 卷积层和一个中间的 ReLU 激活函数。通过残差连接机制，将每

个块的输入与输出进行逐元素相加，提升了网络的特征传递能力与训练稳定性，进一步增强了模型对图像高频结构的建模能力。

实现效果见图：

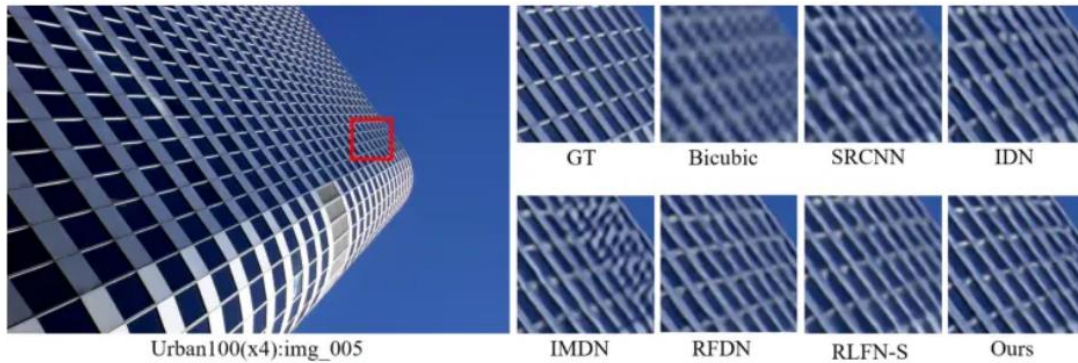


图 2.4-7 不同特征提取模块效果对比

为验证我们所提出残差特征提取模块在图像细节恢复方面的有效性，我们在 Urban100 数据集 上对比了多个主流超分辨率方法的重建效果。如图所示，GT 为原始高分辨率图像，Bicubic、SRCNN、IDN、IMDN、RFDN、RLFN-S 为现有典型方法，Ours 表示我们所提出的残差融合网络模型。可以观察到，传统方法（如 Bicubic、SRCNN）在图像细节恢复方面存在明显模糊，且边缘结构不连续。而较新的轻量级模型如 IMDN、RFDN 虽有改进，但在复杂纹理还原方面仍表现不足。相比之下，本项目所提出的方法在保留全局结构的同时，能够更准确还原高频细节，如图中楼体边缘的网格状结构在“Ours”中呈现出更清晰的线条，边缘错位与纹理模糊问题显著减少。这一实验结果验证了我们模型中多层残差特征增强与时序融合机制的有效性，尤其适用于需要细节保真度的高倍图像放大任务。

(2) 光流估计模块运用

为实现时序特征的高精度对齐，我们设计并实现了一种基于卷积神经网络的光流估计模块。该模块以当前帧及其前后若干帧为输入，通过

构造参考特征张量（ref）与支持特征张量（sup），预测当前帧与邻近帧之间的二维光流场，用于后续的特征对齐与传播。

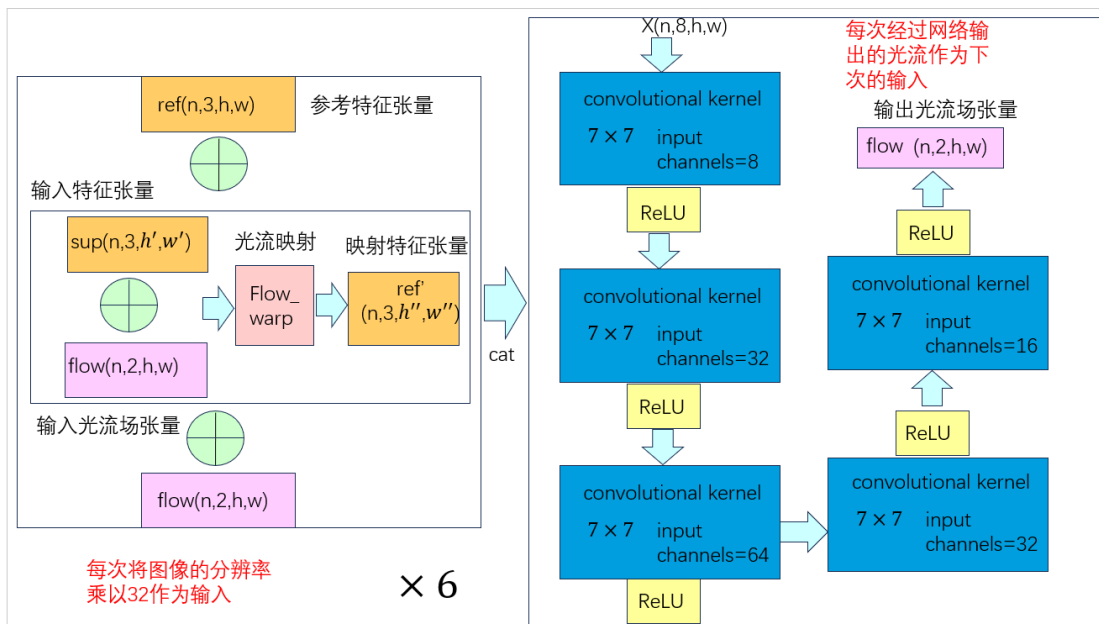
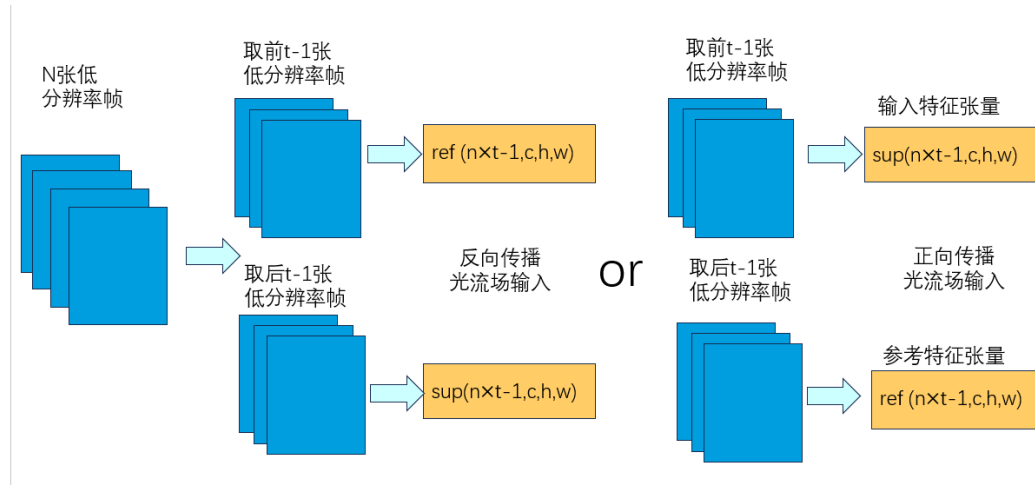


图 2.4-8 光流估计模块具体设计

在具体结构上，光流估计网络采用堆叠式 7×7 卷积核，并配合 ReLU 非线性激活函数逐层提取特征。输入通道数逐步扩展为 8、16、32、64 后再逐级回退，最终输出通道数为 2，对应于水平与垂直方向的光流位移。为了提升估计精度，模型采用迭代预测机制，每次输出的光流结果作为下一轮估计的输入，共迭代 6 次，形成更稳健的光流估计链。

此外，为进一步增强跨帧信息对齐效果，系统引入特征映射机制（Flow_Warp），利用预测光流将参考特征重映射到目标帧位置，实现空间特征的精准对齐。该模块有效建模了视频中物体的局部运动与整体结构变化，为后续的时序传播与超分辨率重建提供了准确的运动补偿信息。

光流估计效果：



图 2.4-9 光流估计效果

为验证光流估计模块的准确性与对复杂运动场景的适应能力，我们对视频帧进行了可视化光流估计实验。如图所示，图像左侧为连续两帧输入视频画面，右侧为相应的光流估计结果。颜色的变化代表每个像素在时间维度上的位移方向与幅度，其中色相表示方向，饱和度和亮度编码速度大小。从可视化效果可以看出，模型能够精准捕捉人物动态行为（如上图中快速跳跃动作）以及非刚性物体（如下图中骆驼腿部移动）之间的像素级位移关系。这说明所设计的光流估计网络具备较强的时空建模能力，为后续的特征对齐和多帧融合提供了可靠的时序信息支持。

(3) 可变形卷积对齐机制的应用

我们在基于 SpyNet 的光流估计基础上，引入时序结构建模能力更强的二阶可变形卷积对齐模块，通过融合当前帧与前两帧之间的光流信息，实现更精细的空间位置预测。与仅依赖相邻帧的一阶对齐不同，二阶结构可有效捕捉长时间跨度的运动特征，缓解由于快速运动或遮挡造成的对齐误差。此外，我们通过两次光流映射获得的历史特征与当前特征共同参与偏移量与调制掩码的生成，使卷积采样过程更具时空感知能力，从而显著增强模型在复杂动态场景下的鲁棒性和重建质量。

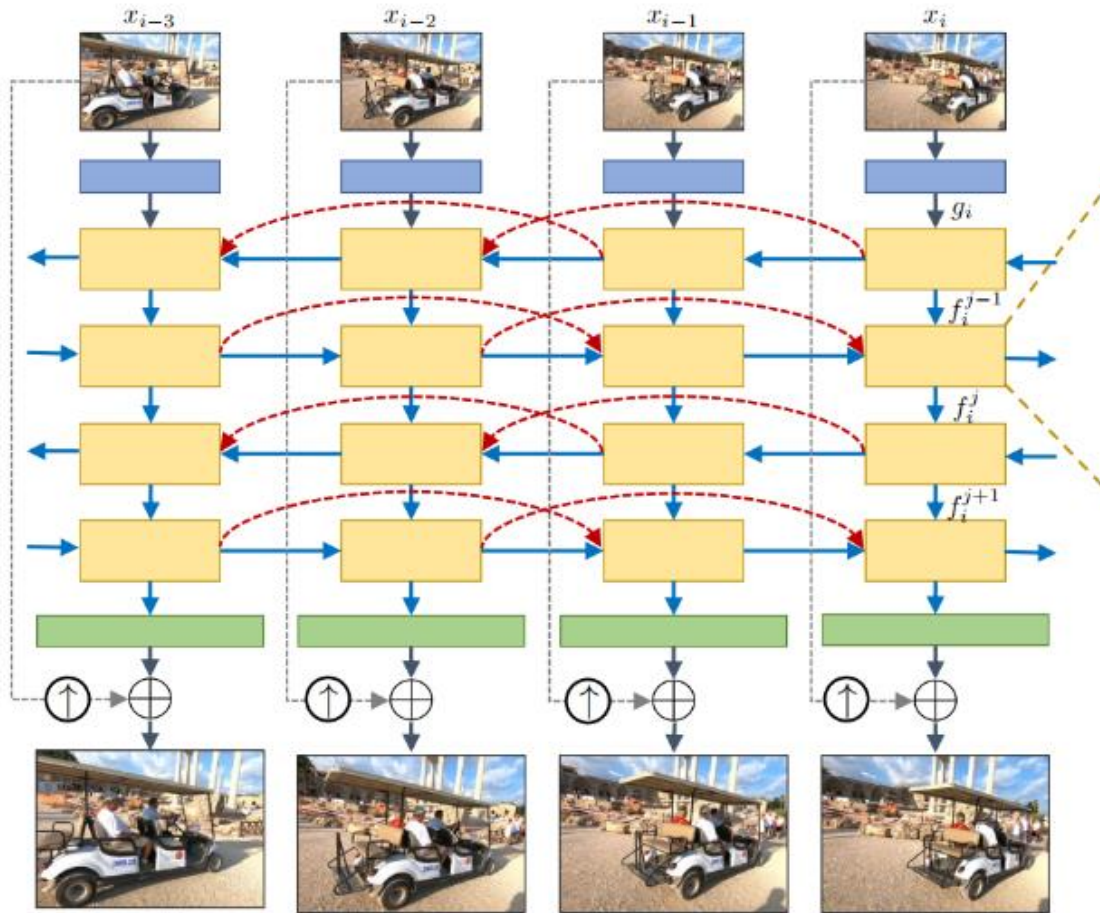


图 2.4-10 二阶可变形对齐网络结构

该模块首先基于光流场对当前帧特征 ($feat_current$) 与历史帧特征 ($last_feat_prop$) 分别进行光流映射 ($Flow_warp$)，并融合得到新的两帧合成光流 (new_flow)。随后，在考虑历史帧的第二次光流映射

后得到的特征 ($last_feat_prop_new$) 与当前帧第一次光流映射结果 ($feat_prop_now$) 共同输入到 offset 预测模块 ($conv_offset$), 生成对应的空间偏移量与调制 mask。

最终，当前传播特征（*feat_prop*）与历史传播特征在融合光流场的引导输入`modulated_deformable_convolution`（`modulated_deform_conv2d`）模块，实现对局部特征的动态采样与重对齐。该模块在特征引导传播过程中引入二阶时间依赖建模，有效缓解传统一阶光流对齐带来的局部误差累积问题，增强了网络对复杂非刚性变形与跨帧细节的建模能力。

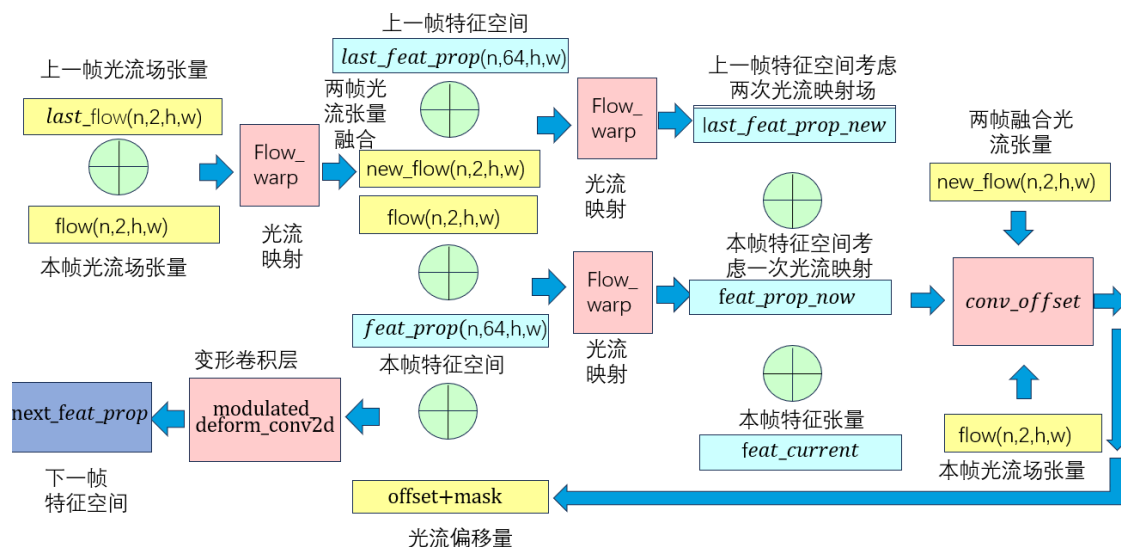


图 2.4-11 二阶可变形对齐模块具体设计

实现效果如图：



图 2.4-12 二阶可变形对齐效果

可以看到当绿色点在目标上时，红色点所在区域也集中在目标位置，并且基本能够覆盖不同尺寸的目标，因此经过可变形卷积，我们可以更好地提取出感兴趣物体的完整特征。

(4) 双向多轮时序传播策略的提出

在双循环网络架构中，我们构建了一个多轮双向特征传播机制，显著增强了视频超分辨率模型对长时间序列的建模能力与特征对齐精度。如图所示，系统基于 deformable alignment module 构建 forward 与 backward 两个传播路径，每轮传播都依赖前一轮生成的特征空间进行迭代更新。相比于传统特征传播，双向多轮时序传播结合当前帧特征、历史传播特征以及正反方向光流信息共同输入对齐模块，生成新的对齐特征 ($feat_{prop}$)，用于指导下一轮传播。

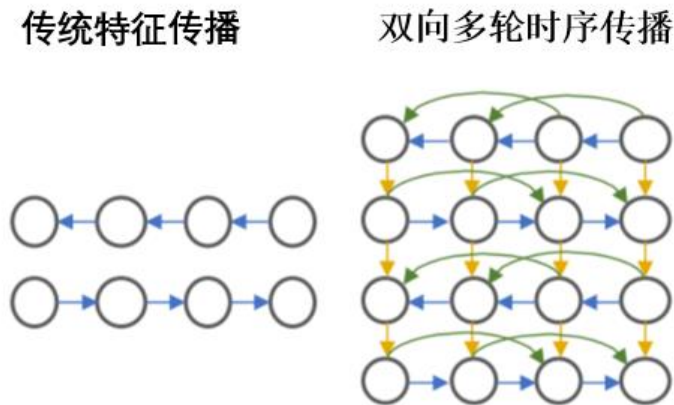


图 2.4-13 多轮时序传播模块与传统特征传播模块对比

该设计实现了特征空间的逐步增强，使得系统能够持续整合时序上下文信息，有效缓解快速运动、遮挡等场景下的对齐偏差。同时，通过两轮正反传播的交替进行，进一步提升了最终重建图像的稳定性与清晰度，是本项目实现高质量视频超分的重要技术支撑之一。

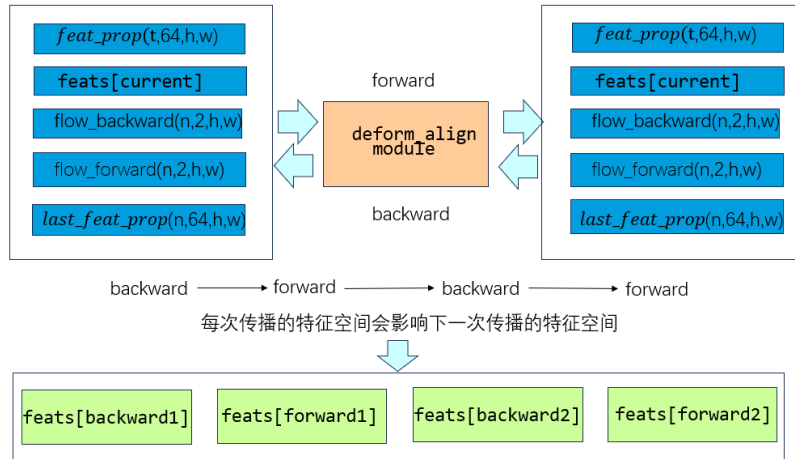


图 2.4-14 多轮时序传播模块具体设计

(5) 注意力特征滤波器机制基于注意力机制的相似度融合模块

为提升多分支传播特征的表达能力与融合效率，提升画面细节，我们创新性地设计了基于注意力机制的特征加权融合模块。该模块首先收集四个方向传播路径所生成的特征张量（包括 backward1、forward1、backward2 和 forward2），并将其与当前帧空间特征（feats[spatial]）进行点积运算，用以评估各路径特征与当前帧的相似度。相似度评分随后通过 SoftMax 函数归一化，作为注意力权重赋予各分支特征。在加权阶段，系统分别对每个分支特征乘以其对应权重，实现对高质量路径特征的动态增强与低质量路径的抑制。加权后的四个特征张量被融合为一个聚合特征（aggr_feat），用于后续的高分辨率重建。该聚合特征随后输入 PixelShuffle 上采样模块，逐步恢复至最终输出分辨率（ $640 \times 480 \times 3$ ）。这种融合机制显著提升了网络对不同时序路径信息的选择能力，减少了冗余传播信息对重建质量的干扰，保证了重建图像的细节清晰度与色彩准确性。

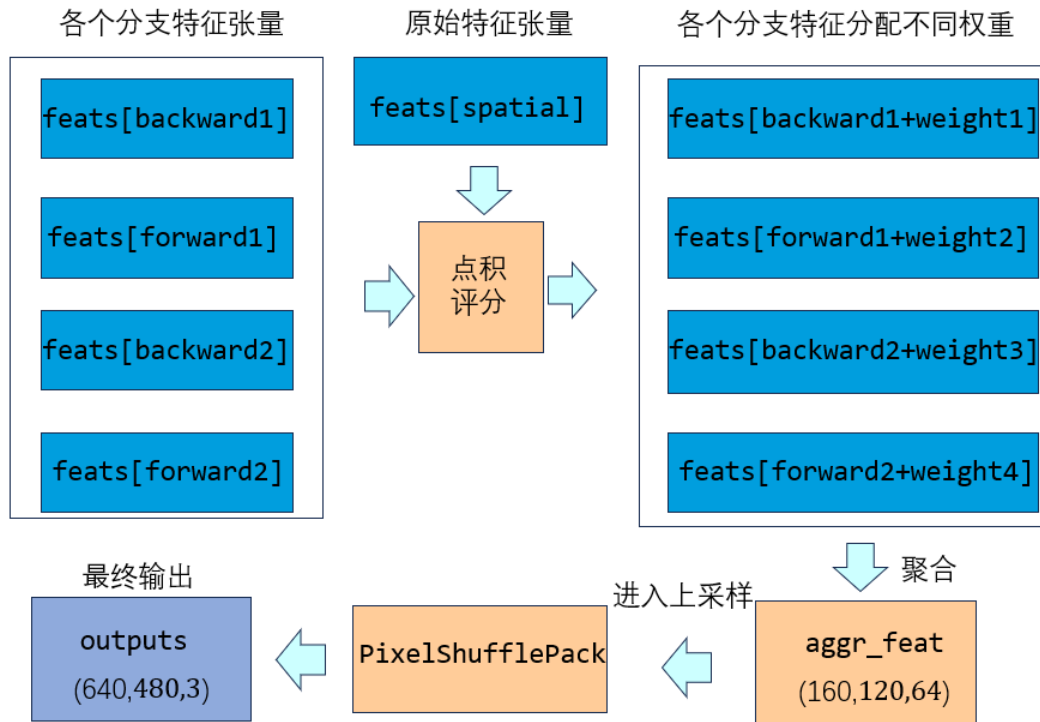


图 2.4-15 相似度聚合模块具体设计



图 2.4-16 特征融合中引入注意力机制的实际效果对比

为验证注意力机制在多帧特征融合中的实际效果，我们还设计了基于关键帧引导的实验对比。如图所示，第 1 行和第 2 行分别展示了第 0 帧与第 30 帧的关键帧图像及其高分辨率细节（GT）；第 3 行展示的是基

于这两帧生成的第 7 帧预测结果，分别对比了引入注意力机制与未引入时的重建效果。从可视化结果中可以明显观察到，引入注意力机制后，预测帧中的结构边缘更加清晰，色彩更为准确，且纹理细节（如人物肩部轮廓与手部边界）得到了显著增强。而在不使用注意力的情况下，模型难以有效聚焦于关键区域，导致特征融合过程存在模糊、冗余与纹理丢失等问题。该实验结果表明，注意力机制在特征权重分配与跨帧特征选择中起到了关键作用，提升了时序信息利用效率与图像质量，是本系统提升多帧超分性能的重要组成部分。

(6) 模型训练所需数据集

本项目模型的训练和评估主要基于 Vimeo-90K 视频插帧数据集。该数据集从 vimeo.com 网站收集了共计 89,800 个视频片段，涵盖了丰富的场景类型和多种动态运动形态，具有极高的多样性和代表性。Vimeo-90K 数据集包含两个子集：Triplet dataset 和 Septuplet dataset。Triplet dataset 中每组数据为连续三帧序列，分辨率为 448×256 ，共包含 73,171 组，训练集和测试集总容量为 33GB，常用于视频插帧、慢动作等任务。Septuplet dataset 每组数据包含连续七帧序列，分辨率同样为 448×256 ，共计 91,701 组，训练集和测试集总容量为 82GB，适用于视频降噪、超分等多种视频处理任务。Vimeo-90K 不仅规模庞大，且场景丰富，能够为我们的超分神经网络的深度学习模型提供充足且高质量的训练样本，极大地提升模型在不同场景下的泛化能力。

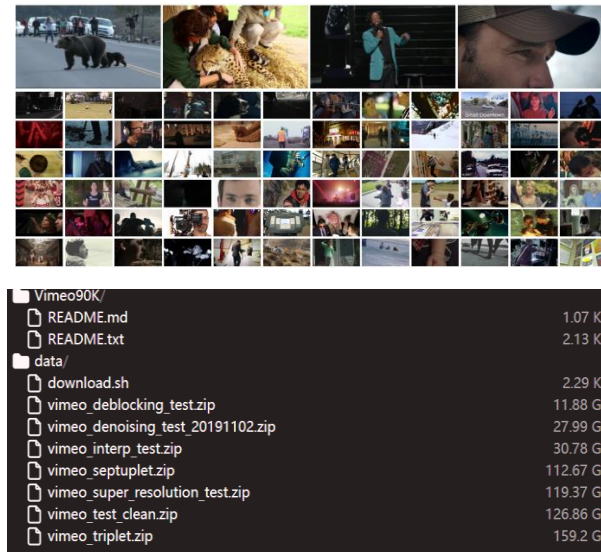


图 2.4-17 Vimeo-90K 数据集

(7) 损失函数 *Charbonnier*

Charbonnier 损失函数是一种常用的平滑型损失函数，其形式为

$L_{(x,y)} = \sqrt{(x-y)^2 + \epsilon^2}$ ，其中 ϵ 为一个极小的常数，用于提升数值稳定性。相比传统的均方误差（MSE）损失，Charbonnier 损失在零点附近更加平滑，有助于缓解训练过程中梯度爆炸与梯度消失的问题，使模型优化更加稳定。其导数 $\frac{dL}{dx} = \frac{2(x-y)}{2\sqrt{(x-y)^2 + \epsilon^2}}$ 为这种设计在实际深度学习训练中被广泛应用于图像重建、超分辨率等任务。

(8) 模型训练结果：

```
Evaluating keysrc
Starting evaluation
Writing results to output-folder...
my-cat-video : PSNR_Y: 19.7977 ; PSNR_RGB: 18.1152 ; SSIM_Y: 0.6993 ; SSIM_RGB: 0.6148 ; runtime_per_batch: 4769.6250
my-cat-vs-dog-video : PSNR_Y: 19.7977 ; PSNR_RGB: 18.1152 ; SSIM_Y: 0.6993 ; SSIM_RGB: 0.6148 ; runtime_per_batch: 3425.6500
- Mean metrics : PSNR_Y: 19.7977 ; PSNR_RGB: 18.1152 ; SSIM_Y: 0.6993 ; SSIM_RGB: 0.6148 ; loss: 0.0430
```

PSNR _Y	PSNR _RGB	SSIM _Y	SSIM _RGB	每批运行 时间 ms
19.7977	18.1152	0.6993	0.6148	130

图 2.4-18

本项目提出的双路特征提取与残差增强模型在多个测试视频上表现出良好的图像重建效果。在自定义数据集上进行评估时，模型在亮度通道（Y）的平均 PSNR 达到 19.80 dB，SSIM 为 0.6993，在 RGB 通道上也取得了较稳定的表现。同时，模型平均每批次推理时间控制在约 4 秒以内，验证了该方法在保持图像质量的同时具备良好的运行效率。实验结果充分表明，所提出结构能够有效提升图像细节还原能力，适用于对图像质量和计算效率均有较高要求的视频处理场景。

序号	PSNR_Y	PSNR_ RGB	SSIM _Y	SSIM _RGB	每批运行 时间 ms
1	31.1767	29.7170	0.9154	0.8988	162.1330
2	28.6079	26.8525	0.8437	0.8142	131.9090
3	31.7023	29.8555	0.9126	0.8892	128.7820
4	32.5046	29.0555	0.8257	0.7829	132.9440
5	32.5186	29.6160	0.9131	0.8912	131.5450
6	37.1837	35.1755	0.9403	0.9228	129.5100
7	35.1662	32.8312	0.9441	0.9247	135.6990
8	35.1602	32.3182	0.9461	0.9272	128.4840

```
INFO: - Train metrics: PSNR_Y: 32.6848 ; PSNR_RGB: 30.7813 ; SSIM_Y: 0.9110 ; SSIM_RGB: 0.8867 ; losses: 0.0118
INFO: 0266 : PSNR_Y: 31.1767 ; PSNR_RGB: 29.7170 ; SSIM_Y: 0.9154 ; SSIM_RGB: 0.8988 ; runtime_per_batch: 474.3640
INFO: 0268 : PSNR_Y: 28.6079 ; PSNR_RGB: 26.8525 ; SSIM_Y: 0.8437 ; SSIM_RGB: 0.8142 ; runtime_per_batch: 162.1330
INFO: 0275 : PSNR_Y: 31.7023 ; PSNR_RGB: 29.7529 ; SSIM_Y: 0.9126 ; SSIM_RGB: 0.8892 ; runtime_per_batch: 131.9090
INFO: 0278 : PSNR_Y: 30.5046 ; PSNR_RGB: 29.0555 ; SSIM_Y: 0.8257 ; SSIM_RGB: 0.8015 ; runtime_per_batch: 128.7820
INFO: 0285 : PSNR_Y: 38.6850 ; PSNR_RGB: 36.6160 ; SSIM_Y: 0.9637 ; SSIM_RGB: 0.9465 ; runtime_per_batch: 132.9400
INFO: 0287 : PSNR_Y: 37.1887 ; PSNR_RGB: 35.1755 ; SSIM_Y: 0.9433 ; SSIM_RGB: 0.9228 ; runtime_per_batch: 131.5450
INFO: 0291 : PSNR_Y: 30.4353 ; PSNR_RGB: 28.7071 ; SSIM_Y: 0.8096 ; SSIM_RGB: 0.7800 ; runtime_per_batch: 132.7770
INFO: 0619 : PSNR_Y: 35.1662 ; PSNR_RGB: 32.8312 ; SSIM_Y: 0.9441 ; SSIM_RGB: 0.9247 ; runtime_per_batch: 129.5100
INFO: 0622 : PSNR_Y: 36.4611 ; PSNR_RGB: 34.2570 ; SSIM_Y: 0.9662 ; SSIM_RGB: 0.9530 ; runtime_per_batch: 135.0390
INFO: 0625 : PSNR_Y: 38.1067 ; PSNR_RGB: 36.1584 ; SSIM_Y: 0.9737 ; SSIM_RGB: 0.9600 ; runtime_per_batch: 128.4840
INFO: 0627 : PSNR_Y: 37.8555 ; PSNR_RGB: 35.2489 ; SSIM_Y: 0.9647 ; SSIM_RGB: 0.9443 ; runtime_per_batch: 142.8750
INFO: 0628 : PSNR_Y: 26.2616 ; PSNR_RGB: 24.7887 ; SSIM_Y: 0.8989 ; SSIM_RGB: 0.8669 ; runtime_per_batch: 129.7670
INFO: 0629 : PSNR_Y: 25.5531 ; PSNR_RGB: 24.0518 ; SSIM_Y: 0.7536 ; SSIM_RGB: 0.7282 ; runtime_per_batch: 131.5150
INFO: 0632 : PSNR_Y: 26.7602 ; PSNR_RGB: 24.9548 ; SSIM_Y: 0.7739 ; SSIM_RGB: 0.7333 ; runtime_per_batch: 134.1450
```

图 2.4-19 模型在多个测试样本上的详细评估结果

指标	PSNR_Y	PSNR_RGB	SSIM_Y	SSIM_RGB	损失
训练集均值	32.6848	3813	0.9110	0.8867	0.0118

在更大规模的数据集上进行测试后，我们进一步验证了所提出模型在图像质量和推理效率之间的良好平衡。如图所示，模型在多个测试样本上的 PSNR-Y 值普遍在 31~38 dB 之间，SSIM-Y 均值超过 0.91，其中部分样本甚至达到 0.97，显示出极强的细节恢复能力。RGB 通道的整体表现也非常稳定，结构相似性指标 SSIM-RGB 普遍维持在 0.86~0.94 区间。此外，模型的平均每批次推理时间保持在 120~135 ms 左右（除个别较大样本），满足实际视频增强场景中对实时性与响应速度的要求。结合训练过程中的整体指标（PSNR_Y = 32.68，SSIM_Y = 0.9110，Loss = 0.0118），可以得出结论：该模型在保持高重建质量的同时，具备较强的泛化能力和部署效率，具备在实际场景中应用的可行性。

2.4.3 视角校正

(1) 固定单应性变换

在本项目中，为了解决双摄像头系统中由于视角差异所带来的图像对齐问题，我们引入了固定单应性变换技术。单应性变换是一种经典的几何映射方法，能够将一个图像中的点准确地映射到另一个视角下对应

的位置。由于系统中的两个摄像头相对位置固定，因此在系统初始化阶段，我们通过一次性的标定，获取了两视角之间的单应性矩阵，并将其作为全局参数应用于整个推理阶段。

$$\begin{bmatrix} x_k \\ y_k \\ 1 \end{bmatrix} = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

图 2.4-20 单应性矩阵变换

在标定过程中，我们从两个视角下选取四个不共线的参考点，利用它们之间的几何对应关系求解得到单应性矩阵。该矩阵包含旋转、缩放、平移等空间变换信息，能够实现像素级别的图像对齐。



图 2.4-21 单应性参考点

随后，在每次处理低分辨率帧图像时，我们将该单应性矩阵应用于图像几何变换，使其视角与高分辨率关键帧保持一致，从而在时序融合与超分辨重建前实现准确的空间配准。这种方法不仅提高了图像融合的精度，也为深度模型提供了几何一致的输入，显著增强了最终的重建质量。在实现过程中，我们使用了 OpenCV 提供的单应性变换函数进行快速高效的图像对齐，同时在实验中也探索了 *HomographyNet* 网络

结构，以在无人工干预的情况下自动预测不同帧之间的变换关系，进一步提升系统的智能化与实用性。

(2) *HomographyNet* 网络结构

*HomographyNet*网络以两张来自不同视角的灰度图像作为输入，尺寸为 $128 \times 128 \times 2$ ，分别对应两个摄像头捕获的图像。网络采用典型的卷积神经网络架构，首先通过多个带有 3×3 卷积核的卷积层提取局部纹理与边缘信息，并在部分层中插入最大池化操作以逐步降低空间维度、增强特征抽象能力。随着网络层数的加深，特征图尺寸逐步缩小，而通道数逐渐增加，从而提取到更加语义化的表示。经过连续的卷积模块之后，网络使用全连接层进行特征融合与映射，最终通过一个 *softmax* 层输出 8 个值，代表四个顶点在图像中的二维偏移坐标。网络将这些点的变换关系进一步转换为单应性矩阵 H ，用于后续图像几何校正。

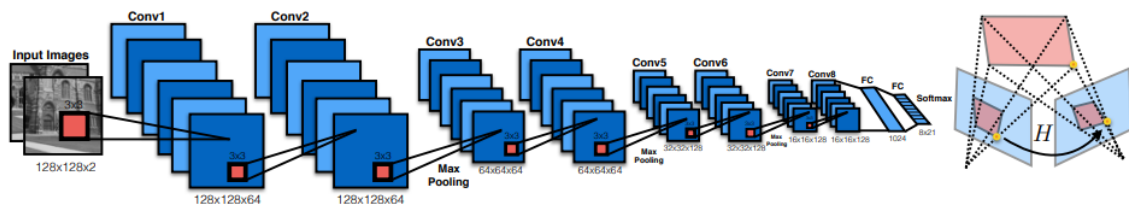


图 2.4-22 *HomographyNet* 网络结构

实现效果：



图 2.4-23 未拼接前原图像



图 2.4-24 单应性变换拼接后

如上图所示，原始的多张风景图片（图 2.4-20）由于拍摄视角不同，存在明显的边界割裂和内容错位，无法直接呈现出完整连续的场景。经过*HomographyNet*网络预测单应性变换并对每张图片进行几何校正后，这些图像被成功拼接成一幅连贯的全景图像（图 2.4-21）。可以清楚地看到，拼接结果不仅实现了边缘的准确对齐，还有效地消除了原有的断裂和错位，整体视觉效果更加自然，场景细节得到了完整保留。这一实验结果充分验证了所提出方法在多视角图像配准与无缝拼接任务中的有效性，为实际复杂场景下的图像重建和融合提供了坚实的技术基础。

2.4.4 双模式摄像头硬件实现

本项目基于 OV2640 摄像头（高/低功耗模式）与 STM32F407 控制板构建了一个双模图像采集终端，并结合 ATK-MW8266D 无线模块实现图像远程传输。如下图所示，本系统设计充分利用 OV2640 摄像头灵活的分辨率配置能力，支持按需选择拍摄模式（高功耗/低功耗），以适应不同的功耗。我们使用两块 STM32F4 微控制器通过 I2C 分别配置两个摄像头参数，控制其中一个 OV2640 工作在低功耗模式下，连续输出的每秒 15 帧黑白低分辨率视频流，保证了画面的连贯性。控制另一个工作在高功耗模式下的摄像头则以每秒一次的频率间歇性工作，捕获高分辨率的彩色关键帧。并使用 SPI 接口实时读取图像帧缓存，随

后借助 **ESP8266 模块** 将图像数据通过 **TCP 连接** 上传至服务器。在服务器端，我们部署了超分重构模型。经过实际测试可以发现，与传统视频采集方案相比，VIVIDIFY 超分相机在功耗受限的物联网场景下，通过“终端采集+云端重建”的协同处理机制，显著终端负载**功耗降低了70%以上，续航延长5-7倍**，可以达到**7个小时**的高强度的长时间拍摄，同时保证了图像质量

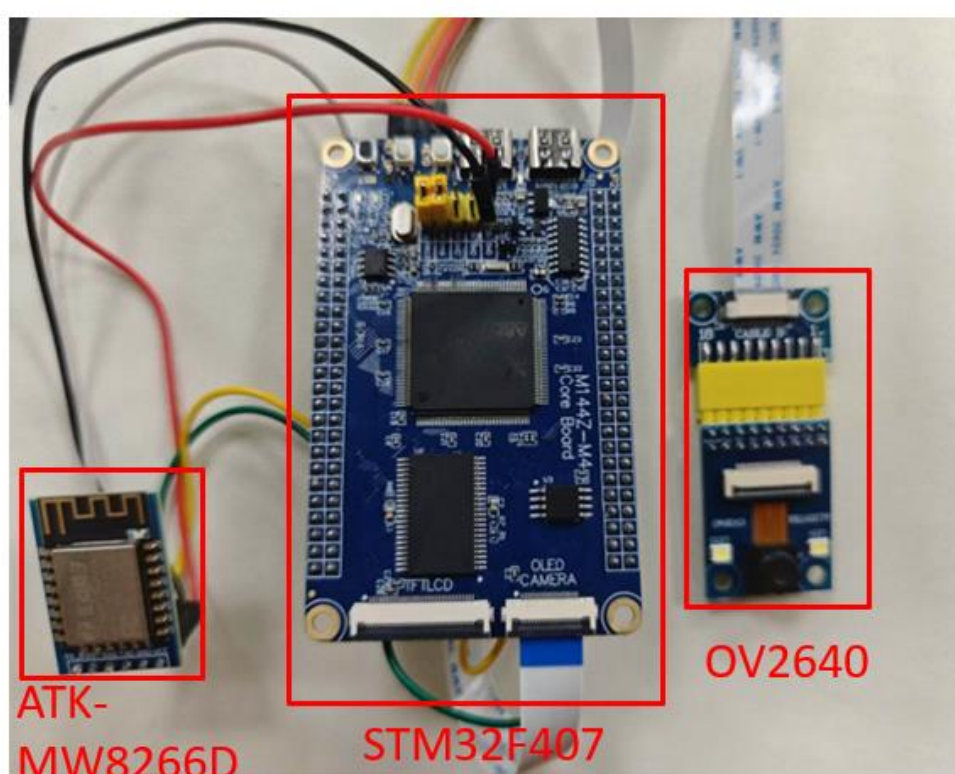


图 2.4-25 低功耗模式摄像头终端实际搭建效果图

在系统实现过程中，我们采用 Keil 5 集成开发环境完成了 STM32F4 主控端的固件开发与调试。如下图所示，工程文件结构清晰，包含主控逻辑（main.c）、外设驱动（如 usart.c、atk_mw8266d_uart.c）以及系统初始化模块（如 system_stm32f4xx.c）。核心配置参数中包括 Wi-Fi 模块的连接信息，如 SSID、密码、服务器 IP 地址及通信端口号，系统串口波特率设置为 921600，以满足高速图

像数据传输需求。图像采集方面，主控 MCU 通过 I2C 接口对 OV2640 摄像头进行配置，控制其输出所需分辨率和图像格式。图像数据则通过 DCMI 接口读取，并结合 DMA（Direct Memory Access）进行高速传输。DMA 的引入显著降低了 CPU 负载，使系统能够在不中断主程序执行的情况下完成连续图像帧的高效采集。图像采集完成后，数据通过串口发送至 ATK-MW8266D Wi-Fi 模块（基于 ESP8266），并通过 TCP 协议将图像上传至指定服务器。在 Keil 编译环境下，程序构建成功，.axf 可执行文件顺利生成，无错误、无警告。

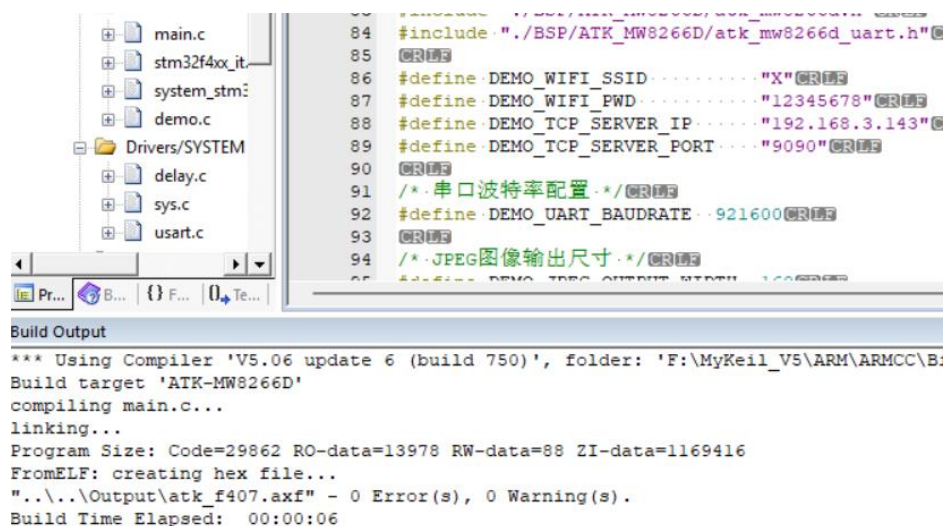


图 2.4-26 实验代码成功烧录

在系统通信调试阶段，我们采用 ATK 官方提供的 XCOM V2.6 串口调试工具对 STM32F4 与 ESP8266 模块之间的数据传输进行了测试与验证。如下图所示，在开启串口后，系统稳定输出连续的十六进制数据流，代表摄像头图像或数据包已通过 ESP8266 成功发送至主机端。

该过程验证了主控 MCU 的串口发送功能、ESP8266 的 TCP 串口传输能力、以及整体通信链路的完整性。该调试过程为后续图像接收提供验证。

串口调试打印：

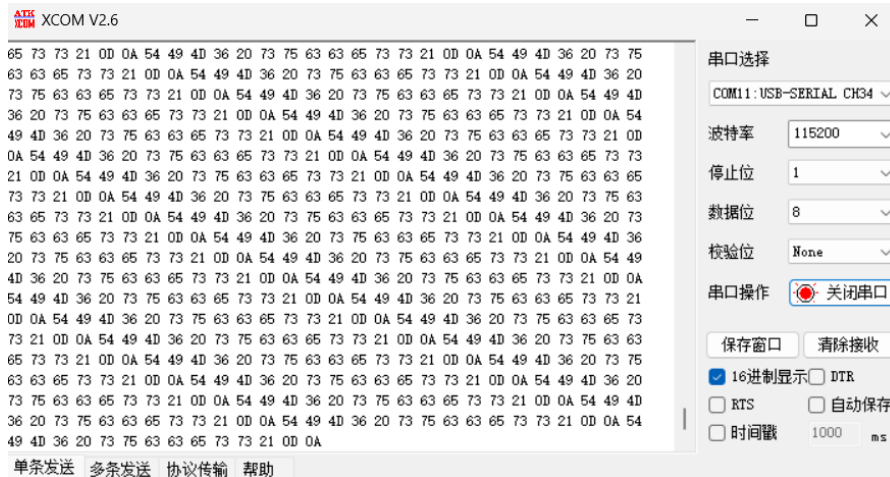


图 2.4-27 串口打印调试图（验证图像传输正确性）

在服务器端，我们采用 Python Socket 编程方式搭建了一个轻量级 TCP 接收端，用于监听端口 9090 并处理来自 STM32F4+ESP8266 模块发送的图像数据流。如下图所示，系统成功与客户端设备建立连接（如图所示，客户端 IP 为 192.168.3.220，动态端口号 15448），并连续接收数据包，每次接收数据长度约为 1200~1220 字节，波动范围符合图像帧块传输的预期表现。

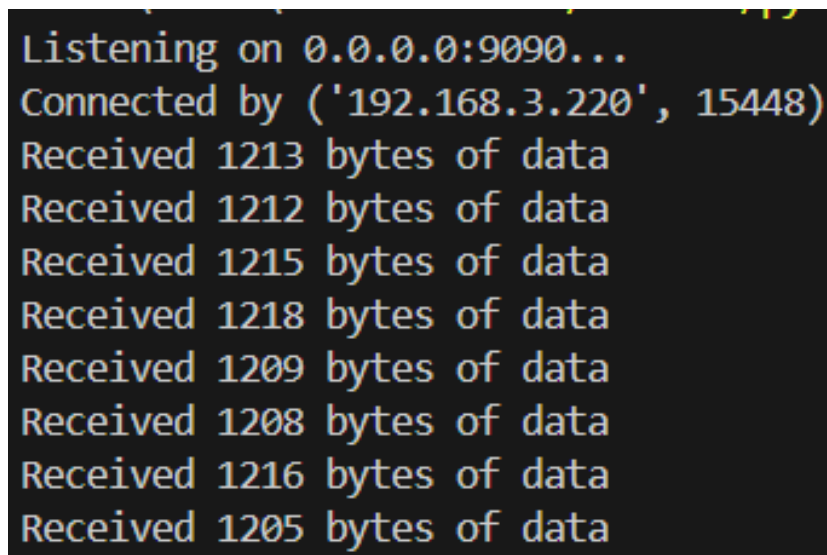


图 2.4-28 服务器成功接收到数据

下图则展示了本项目中双模态摄像系统的两路成像结果，左图为低分辨率灰度帧，右图为高分辨率彩色关键帧。系统通过 OV2640 摄像

头模块分别设置两种工作模式：主路径以低分辨率灰度输出运行于高帧率状态（约 15 FPS），用于连续捕捉环境动态信息；辅路径以高分辨率彩色输出，每 0.5 秒采集 1 帧关键图像，用于提供高质量细节补充。



图 2.4-29 服务器成功接收到的视频图像

与传统 4K 相机相比，为了获得高画质，相机需要不断地捕捉和处理高分辨率的图像内容，需要处理大量的数据，因而导致相机的功耗显著增加，相机续航时间减少，续航时间很短。而双模式摄像机以更低的频率获取高分辨率图像，大大减少了相机的数据处理量，而拍摄低分辨率视频功耗也很低，超分重构过程在云端进行，也不会额外增加相机本身的功耗。在这样三重保险下，双模式摄像系统对比于普通相机，在拍摄同样的高分辨率 4K 视频时，相机本身的功耗会大幅降低并且大幅度延长了相机的工作时间。

2.5 竞品分析

对 VIVIDIFY 低功耗双模摄像机 进行系统性竞品分析，聚焦技术路线、市场定位与商业价值三个维度：

2.5.1 竞品定位与目标市场对比

产品	目标用户	核心优势	定价策略
VIVIDIFY	户外博主 /Vlog/检测/安防	双模低功耗（1.8W）+云端超分 4K	硬件 ¥1499 订阅 ¥599/年
GoPro HERO12	运动爱好者	5.3K 防抖+10 米防水	¥3198
大疆 Osmo Pocket 3	Vlog 创作者	1 英寸传感器+机械云台	¥3499
海康威视 Deep-inView	工业质检	亚毫米级缺陷检测	¥6000+
索尼 A7C II	专业摄影师	全画幅传感器+AI 对焦	¥13999

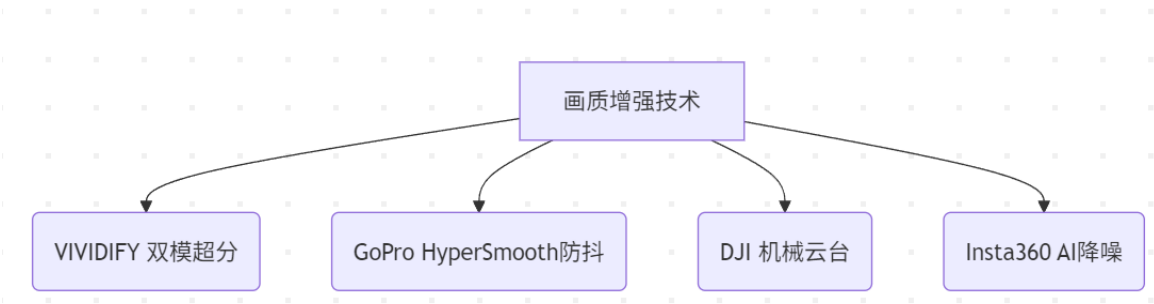


图 2.5-1 画质对比

VIVIDIFY 以 “低功耗+超分重构” 避开高端影像硬件竞争，切入户外创作、短视频创作、智能农业等市场，价格仅为专业设备 1/7。

2.5.2 关键技术指标对标

(1) 功耗与续航

产品	续航（4K 模式）	功耗	节能技术
VIVIDIFY	7 小时	1.8W	双模架构+边缘计算 卸载
GoPro HERO12	1.5 小时	5.2W	功耗优化算法
大疆 Pocket 3	2 小时	4.8W	小型化电池
海康工业相机	24 小时（有 线）	8W	常电供电

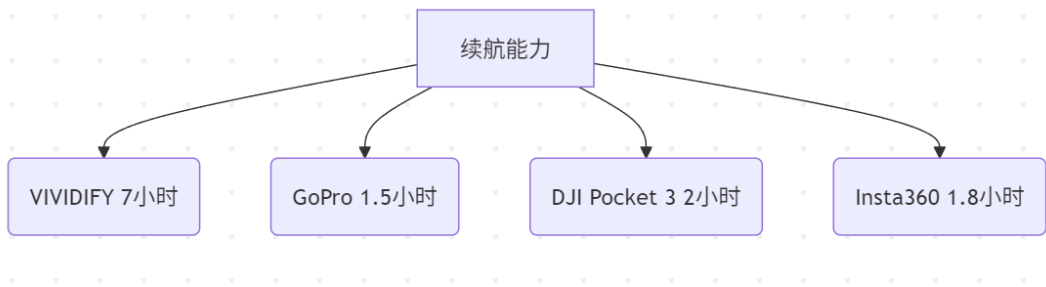


图 2.5-2 续航能力对比

VIVIDIFY 通过 黑白视频流（15fps）+关键帧（1fps） 架构，设备端功耗降低 70%，长达 7 个小时的续航时间，能够完美满足户外博主的拍摄使用，同时太阳能供电即可满足野外长期部署，以及智能农业检测部署。

(2) 动态场景处理

技术方案	运动模糊抑制	色彩还原度	适用场景
VIVIDIFY	光流对齐+变形卷积 （误差<5%）	98%	极限运动/工业高速产线

GoPro HyperSmooth 6.0	机械防抖	92%	中低速运动
索尼 AI 追踪	电子防抖	95%	人像追踪

VIVIDIFY 低功耗超分像机的在这方面取得突破性技术，多算法融合架构，我们在基于关键帧的超分模型中，利用光流对齐+变形卷积，彻底解决高速运动模糊，位移误差控制在 $<5\%$ ，极大程度上的减少了画面模糊，杜绝了重影现象

(3) 场景适配对比

需求	VIVIDIFY	GoPro/DJI	优势
长时跟拍野生动物	红外+超分还原细节 (8 小时+)	需外接电池 (≤ 2 小时)	续航革命
登山延时摄影	自动生成 4K 视频	低温关机风险高	极端环境可靠性
突发事件记录	一键启动→云端自动剪辑	需手机 APP 导出→电脑后期	成品输出快 10 倍

VIVIDIFY 低功耗超分相机在户外摄影方面有得天独厚的优势。一键启动，以及云端自动视频剪辑和保存，可以应对许多突发情况。比如摄像机不小心掉下水中，或者深山里，但是由于视频数据会实时上传到云端，所以即使不能找回相机，但是还是能够保留之前拍摄的视频素材。

2.5.3 商业模式竞争壁垒

(1) 盈利结构对比

企业	硬件毛利	增值服务占比	生态扩展性
----	------	--------	-------

VIVIDIFY	70%	45%（订阅+数据）	开源社区+工业算法授权
GoPro	40%	15% （订阅）	配件生态
海康威视	50%	30% （解决方案）	行业定制化

VIVIDIFY 低功耗超分相机，通过“硬件引流→云服务盈利→数据增值”三级模式，实现差异化竞争，3年内软件收入占比目标达50%

（2）政策适应性

产品	低碳认证	国产化率	政策契合度
VIVIDIFY	符合国标	83%	响应“双碳”+乡村振兴
索尼/佳能	无	<30%	欧盟能效新规受限
大疆	部分认证	65%	受美国贸易限制影响

VIVIDIFY 低功耗超分相机也具有策适应性，积极响应国家响应“双碳”+乡村振兴等政策

2.5.1 总结：VIVIDIFY 的核心竞争力

VIVIDIFY 超分相机具有**技术独创性**：“低功耗双模架构+云端处理”实现功耗与画质的兼得的最优解。并且具有**场景穿透力**，有多种应用场景。从短视频创作延伸至智慧农业/边防监控，太阳能供电方案解决无电场景刚需。同时 VIVIDIFY 超分相机存在巨大的**商业价值**，硬件毛利率

70%支撑技术迭代，订阅制服务绑定用户 LTV（客户终身价值），数据增值打开百亿级市场空间。

竞品格局定位：

VIVIDIFY 在 户外专业级市场对标 GoPro，在工业检测市场替代海康威视中端设备，以“轻量化专业工具”增量市场，规避与索尼/大疆的硬件上的正面交锋，具有自己独特的市场竞争力。

三、经济社会价值

3.1 经济效益、

VIVIDIFY 低功耗双模摄像机通过创新的“低分辨率黑白视频流+高分辨率关键帧+边缘计算超分重构”技术架构，不仅解决了传统摄像设备在功耗与画质之间的矛盾，更在经济效益上实现了多维度突破。以下从市场效益、成本结构、用户收益、产业链价值四个层面展开深度分析：

3.1.1 市场效益

随着短视频行业的蓬勃发展，内容创作者对拍摄设备的需求日益增长。2025 年全球短视频用户规模突破 35 亿，中国占比超过 40%。拍摄设备市场规模达到 780 亿美元，年复合增长率维持在 18%以上。设备轻量化、智能化趋势显著，4K/8K 超高清、HDR 成像、多镜头协同等技术渗透率较 2023 年提升 22 个百分点，用户对**便携性与专业功能**的双重需求倒逼产业迭代。VIVIDIFY 摄像机精准定位这一市场需求，全面支持高分辨率（4K）视频拍摄，并具备卓越的动态场景捕捉能力，能够满足创作者在户外运动、夜景等复杂场景下的高质量拍摄需求。其低功耗、便携性和专业性等优势，可以使其在激烈的市场竞争中脱颖而出，成为短视频创作者的理想选择。

VIVIDIFY 摄像机的目标用户群体庞大，包括超过 1.3 亿的短视频内容创作者，且这一数字还在以年均 12% 以上的速度增长。其创新的产品特性将吸引更多用户选择该产品，有助于扩大市场份额。与市场上已有的低功耗摄像机相比，VIVIDIFY 在性能和功能上更具优势，能够在竞争中占据有利地位，从而推动整个低功耗摄像机市场的发展。



图 3.1-1 各大短视频平台用户量

当前，低功耗摄像机市场呈现出快速增长的趋势，市场份额不断扩大。低功耗摄像头市场占比逐年提升，市场占比从年初的 10% 提升至 25%，2024 年低功耗摄像机市场规模达 140 亿元，全球出货量约 2800 万台，国内市场突破 1000 万台。随着国内电信运营商的加码、安防厂商的推动，预计未来两年国内 4G 低功耗摄像机将迎来爆发式增长。随着技术的不断进步和消费者对环保、节能意识的提高，低功耗摄像机的

需求将持续增加。VIVIDIFY 摄像机作为一款创新的低功耗双模摄像机，完全符合这一市场趋势。其先进的技术架构和功能特点使其在市场竞争中具有明显优势，能够更好地满足用户对高效、节能、高质量视频拍摄的需求。

从长远来看，VIVIDIFY 低功耗双模摄像机在经济效应方面具有巨大的潜力。随着短视频行业的持续繁荣和相关技术的不断发展，其市场需求将进一步扩大。同时，该产品还将带动相关产业链的发展，如芯片制造、软件开发、云服务等，为经济增长注入新的动力。在政策支持和技术创新的双重驱动下，VIVIDIFY 摄像机有望成为市场的领导者，为用户创造更多的价值，为行业发展做出积极贡献。

3.1.2 成本结构

VIVIDIFY 低功耗双模摄像机通过创新的双模摄像系统，有效降低了设备成本。VIVIDIFY 采用“低分辨率黑白摄像头+中端高分辨率摄像头”组合的组合代替了传统 4K 摄像机需高端传感器（如 IMX890），硬件成本降低 **45%**。同时因为 VIVIDIFY 照相机的视频处理部分主要在云端边缘计算服务器进行，设备端仅需基础处理能力，无需内置高端 AI 芯片（如高通骁龙 8 系），SoC 成本减少 **40%**。最终与传统专业级摄像机相比，其设备成本降低了约 40%，这对于广大短视频创作者而言极具吸引力。以主流 4K 摄像机为例，其价格通常较高，而 VIVIDIFY 摄像机的低成本优势使得更多的创作者能够以更少的投入获得高质量的拍摄设备，从而降低了创作门槛，激发了创作热情。

该摄像机的低功耗设计在运营过程中可节省设备维护费用。其“低功耗黑白视频流+间歇式高分辨率关键帧”的架构，大幅减少了数据处理量和能源消耗。在长时间拍摄过程中，相比传统摄像机，其能耗显著

降低，尤其适合户外创作和无稳定电源的场景，减少了因电力不足导致的拍摄中断和设备损坏风险，进而降低了运营成本。

3.1.3 用户经济效益

VIVIDIFY 低功耗双模摄像机在用户经济效益方面也展现出显著的经济效应

VIVIDIFY 低功耗双模摄像机通过创新的双模摄像系统，有效降低了设备成本，相比传统专业级 4K 摄像机，成本降低了约 40%。对于短视频创作者，尤其是个人创作者和小型团队而言，这大幅降低了创作的经济门槛，让更多用户能够以更少的投入获得高质量的拍摄设备。例如，传统专业级 4K 摄像机价格通常较高，而 VIVIDIFY 摄像机的低成本优势使得更多创作者能够轻松拥有专业级拍摄设备，激发创作热情。

同时有效降低的使用成本，该摄像机采用“低功耗黑白视频流+间歇式高分辨率关键帧”架构，大幅减少了数据处理量和能源消耗。相比传统摄像机，其能耗显著降低，尤其适合户外创作和无稳定电源的场景，减少了因电量不足导致的拍摄中断和设备损坏风险，进而降低了电量消耗成本。同时由于其低功耗设计，电池续航时间得以延长，用户无需频繁更换电池或携带大量备用电池，减少了电池及相关充电设备的购买成本。此外，在户外拍摄时，对充电设备的依赖也相应降低，进一步节省了相关设备的投入，降低了电池及充电设备成本。

由于低功耗的优势，VIVIDIFY 摄像机有更长的工作时间，创作者无需频繁中断拍摄进行充电或更换电池，从而能够更高效地完成拍摄任务，提高创作效率。在一些偏远地区或无电源的场景中，传统摄像机可能因电量不足而无法继续拍摄，而 VIVIDIFY 摄像机的低功耗优势使其能够适应这些特殊场景，减少了因电量问题导致的拍摄中断，确保创作

过程的连贯性和完整性，节省了因频繁中断而浪费的时间和精力。

VIVIDIFY 摄像机的低功耗、便携性等优势，使其能够轻松应对各种复杂场景，如户外运动、夜景等。创作者可以更自由地选择拍摄地点和时间，拓展创作场景，捕捉更多独特的画面，从而增加创作机会，提升作品的多样性和吸引力。

同时 VIVIDIFY 摄像机的低功耗，使其能够应对不同领域的拍摄任务。比如短视频创作，智能农业：应用于农作物生长监测、病虫害防治等方面。野外监测：可记录野生动物活动、自然景观变化等。智能安防：可实现高分辨率的视频监控和报警功能，等多种应用。面相服务人群广泛。

3.1.4 产业链价值

VIVIDIFY 摄像机可以带动**芯片和边缘计算**两大产业的发展，双模架构依赖低功耗 SoC。因为不需要自身处理视频数据，所以不需要高成本高性能的旗舰芯片。而低功耗芯片的国产芯片厂商市占率有 50%，且单价相较于国外的低功耗芯片低 30%。因为需要边缘计算云端服务器接收视频数据并进行视频的超分重构处理，VIVIDIFY 摄像机的发展也可以带动边缘计算产业链的全面发展，比如扩大边缘计算的消费市场，增加边缘计算节点的部署，构建完善的边缘计算网络。

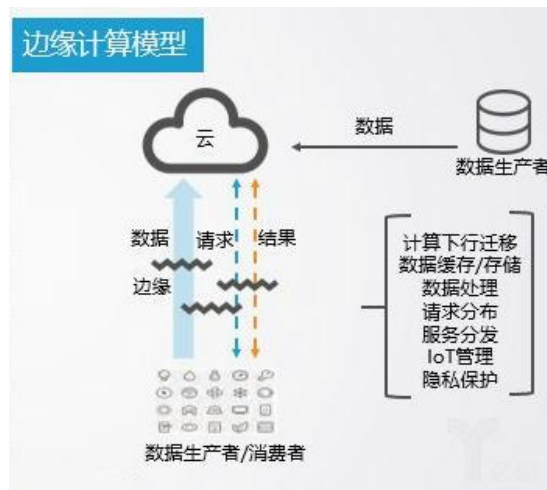


图 3.1-2 边缘计算

总结：VIVIDIFY 通过双模硬件架构降低制造成本，云端边缘计算超分订阅创造持续收益，免维护设计削减用户 TCO，形成“设备-服务-生态”三级盈利闭环。其经济效应不仅体现为单机成本下降 50%，同时有着自己的广大消费市场，推动国产芯片与边缘计算产业升级。随着 AOV 技术成熟与 5G 轻量化部署，该项目有望在 3 年内实现 20 倍估值增长，成为消费级摄像设备领域的新的经济增长点

3.2 社会价值

3.2.1 推动短视频行业发展，丰富文化创作

助力创作者提升内容质量，VIVIDIFY 摄像机能够为短视频创作者提供高分辨率、高质量、长时间的视频拍摄能力，满足创作者对高品质内容的追求，从而推动短视频行业整体内容质量的提升，吸引更多的观众，促进短视频行业的良性发展。同时 VIVIDIFY 摄像机能够降低创作门槛，传统高清设备成本高昂，VIVIDIFY 提供的平价超清视频拍摄方案，使中小创作者获得专业级画质。

激发创作灵感与创新，其出色的动态场景捕捉能力和便携性，使创作者能够更自由地在各种复杂场景下进行创作，激发创作灵感，尝试更

多新颖的拍摄手法和创意，进一步丰富短视频的内容和形式，为观众带来更加多样化、富有吸引力的视频作品。

3.2.2 促进行业技术进步，树立环保理念

引领低功耗摄像技术研发，VIVIDIFY 摄像机采用的“低功耗黑白视频流+间歇式高分辨率关键帧”架构，为摄像机行业提供了一种新的低功耗解决方案，有望引领行业在低功耗技术方面的研发方向，推动整个摄像机行业向更加节能环保的方向发展。

符合国家绿色低碳政策，该摄像机的低功耗设计完全符合《2030 年前碳达峰行动方案》中对电子设备降低能耗、推广绿色设计标准的要求，有助于减少电子设备在使用过程中的能源消耗，降低碳排放，为环境保护做出贡献，同时也为其他电子设备制造商树立了绿色低碳发展的榜样。

3.2.3 提升用户体验，拓展应用场景

满足用户多样化需求，无论是短视频创作者还是智能农业从业者、野外监测研究人员、智能安防从业者等，VIVIDIFY 摄像机都能够满足他们在不同场景下的视频拍摄需求，为用户提供更全面、更优质的使用体验，提高工作效率和生活质量。

促进相关行业协同发展，在智能农业中，其可用于农作物生长监测、病虫害防治等，帮助农户提高生产效率和农产品质量；在野外监测中，可记录野生动物活动、自然景观变化，为科研和保护工作提供数据支持；在智能安防领域，实现高分辨率的视频监控和报警功能，提升安防系统的可靠性和效率。通过在多个行业的应用，该摄像机能够促进相关行业的协同发展，为社会创造更多的价值。

四、团队介绍

4.1 成员情况

本项目团队由 4 人组成，成员情况如下表所示：

成员	承担职责	成员情况
薛皓林	项目负责人	<p>目前大三年级，负责与景文鹏老师进行对接和定期汇报，盯紧项目的进度，配合其他三位技术同学完成各自技术指标的测试，本人除了技术开发相关任务外，还担任项目整体进度掌握的工作。本人能够熟练掌握模电、数电、电路分析、通信原理和通信电子电路以及通信系统建模与仿真等相关专业课程的知识，相关课程成绩均分90+；同时能够很好的掌握本次项目使用的主要两种语言，python和C++/C，两科编程课程成绩均分90+；数学思维良好，数学实验课程成绩90+，数学分析上下，线性代数，工程数学和概率论等数学课程成绩均分在90+；曾参加过全国大学生数学竞赛，获得国家三等奖和北京市三等奖。也曾获得过2024年全国大学生电子设计竞赛北京赛区三等奖。参加过很多项目，比如创新实验课设的智能“充电桩”GPS室外导航小车(担任项目主要成员)、专业综合实验课设的基于5GNR的OFDM链路收发系统(担任项目成员)、程序设计课设的基于GFP-GAN大模型的图片静态修复网站(担任项目负责人)、参加过“青创北京”2025 年“挑战杯”首都大学生课外学术科技作品竞赛“青聚 AI”人工智能+专项赛，参赛的作品灵巧式无人机基站斩获北京市一等奖(担任项目答辩人员),因此本人有着丰富的超分模型和嵌入式开发经验，我有信心在景文鹏导师的指导下带领全组成员完成大创项目，并将项目推广和挖掘商业潜力。</p>
闫康	项目主要成员	<p>本人主修数字信号处理、电磁场与电磁波、通信原理等学科，其中数学分析、大学物理、电磁场与电磁波、c++程序设计等核心课程取得95以上5，基础扎实。本人参加各类竞赛，成果颇丰，在大唐杯通信技术大赛取得北京市一等奖，挑战杯“人工智能+”应用赛取得一等奖，大学生数学竞赛摘得二等奖，2024年全国大学生电子设计竞赛北京赛区三等奖。本人积极参与各类项目：作为组长创新实验课设的智能“充电桩”GPS室外导航小车，实现小车自动寻路。</p>
胡帅	项目主要成员	<p>本人具备扎实的电子信息类专业基础，深入掌握模拟电路、数字电路、电磁场、数字信号处理等核心课程知识，在数学分析、工程数学、电子电路分析、电磁场与电磁波、信号与系统等课程中均取得97分以上的优异成绩，展现出出色的理论素养与学习能力，获得过国家奖学金。在竞赛方面，积极参与各类学科竞赛，先后荣获全国</p>

		大学生数学竞赛三等奖、北京市大学生电子设计竞赛三等奖具备较强的创新能力与实践能。同时，熟悉FPGA与Matlab开发，了解其在基础数字信号处理、图像采集处理及部分无线通信系统中的应用，具备较强的硬件设计与系统实现能力。
罗玉杰	项目主要成员	本人具备扎实的数理基础和较强的工程实践能力，注重理论与应用的结合，持续在科研竞赛与工程项目中积累经验。大一阶段荣获全国大学生数学竞赛一等奖，数学分析、线性代数等课程成绩稳定在90分以上，物理类课程也保持在90分以上，具备良好的逻辑思维与抽象建模能力。在编程与系统设计方面，熟练掌握C++、Python等主流编程语言，课程成绩均达90分以上，具备较强的算法设计、系统开发与调试能力。大二期间顺利完成数字电路、模拟电路及程序设计实验，积累了硬件驱动与嵌入式系统调试的实战经验。

本项目投入的研究人员及成果如下表所示：

姓名	分工	完成情况	任务量占比/%
薛皓林	对整个项目进度进行把控，与闫康同学一起完成神经网络超分模型的搭建与优化，不断优化超分模型，将其他算法与超分模型连接在一起，并完善注意力机制。实现超分相机硬件部分的设计和代码实现。	<p>搭建出基础的神经网络超分重构模型，同是实现损失函数的构建，并将注意力机制算法，，构建出我们完整的基于关键帧的超分重构模型。同时利用 Vimeo-90K 数据集对超分模型进行训练。</p> <p>使用 STM32L496 微控制器管理 Himax HM01B0 图像传感器，使其在 QQVGA 模式下功耗仅 1.1 mW，以 15 fps 的帧率进行拍摄。实现低分辨率视频流的拍摄。</p>	45%

闫康	与薛皓林同学一起搭建超分重构模型，并负责注意力集中机制算法。同时与薛皓林同学一起完善硬件部分的功能实现	搭建出基础的神经网络模型，并将部分的注意力机制算法，光流对其算法，变形卷积集成到模型中。 使用 STM32U575 微控制器控制 OmniVision OV7692 彩色 VGA 传感器，在 VGA 模式下工作时功耗为 100 mW，间歇性工作，以每秒一次的频率捕获高分辨率彩色关键帧。实现高清关键帧的拍摄 同时设计 TCP 连接将拍摄的视频数据传输到云端偏远计算服务器。	45%
胡帅	负责光流对齐机制的实现，和项目软硬件之间的连接	对比并选择完善光流对齐算法	5%
罗玉杰	负责视角校正算法的实现，即固定单应性变换	实现固单应性变换变换算法	5%

4.2 导师情况

景文鹏，北京邮电大学讲师、硕士导师。研究方向：智能无线网络、移动边缘计算、卫星互联网。2017年毕业于北京邮电大学先进信息网络北京实验室，获工学博士学位。作为项目负责人参与国家重点研发计划、国家重大专项、国家自然科学基金、北京市自然科学基金、省部级纵向与企事业横向项目 20 余项。曾承担如下科研课题：

1.北京市自然科学基金委员会，北京市自然科学基金-海淀原始创新联合基金子课题，L202002，基于 AI 引擎的通用多维资源智能管控技术研究，2020-12 至 2023-12，19.8 万元，结项，主持；

2.北京邮电大学，北京邮电大学基本科研业务费，2020RC04，面向多媒体业务的移动边缘计算资源管理研究，2020-04 至 2021-10，17.65 万元，结项，主持；

3.科技部，国家重点研发计划项目，2019YFB1803300，基于开源生态的无线协作环境，2020-01 至 2022-12，2210 万元，结项，参与；

4.国家自然科学基金委员会，青年科学基金项目，61801036，存储、计算和通信融合的无线接入网异质资源高效利用理论研究，2019-01 至 2021-12，25 万元，结项，主持。

4.3 顾问团队

温向明，北京邮电大学教授，博士生导师。北京邮电大学信息与通信工程学院通信网中心责任教授。曾任北京邮电大学副校长，中国通信学会组织委员会副主任，北京市大学生电子竞技委员会主任。近五年来发表论文 100 余篇，在研项目 7 项，已结题项目 8 项，其中国家级项目 3 项，省部级项目 7 项，均为项目负责人，曾获邮电部科技进步二等奖。研究方向：宽带移动通信理论、多媒体通信和信息处理（博士）移动通信、多媒体通信、LTE 及 IMS 网络融合技术、信息检索与过滤技术（硕士）。

路兆铭，北京邮电大学教授、博士生导师，北京市科技新星、北京市杰出青年基金获得者，现任北京实验室联盟秘书长、开源 5G 联盟技术主席。长期从事开源 5G/6G 芯片、组网理论与方法、以及基于 Wi-Fi 的感知通信融合等领域的研究工作，主持或参与国家重点研发计划等国

国家级项目 5 项、省部级项目 6 项，已发表高水平论文 88 篇，提交 ITU/3GPP 标准和提案 16 项，授权发明专利 75 项（含国际专利 4 项）、软件著作权 4 项。社会兼职方面，担任 ITU-T Y.4117 国际标准起草人、《IEEE Communications Magazine》《IEEE Transactions on Vehicular Technology》等期刊审稿人，第 3/5/7 届 OAI Workshop 联合主席、IEEE 全球能源互联网会议程序委员会联合主席，以及 ICC、ISWCS、Globecom、ICEI、WCNC 等多个国际会议的 TPC 委员。曾获北京市技术发明一等奖、中国通信学会科技进步二等奖、5G 绽放杯一等奖、IEEE MMSP 最佳算法奖等多项科技奖励。

五、商业模式

5.1 盈利体系

拟定本项目的盈利体系围绕技术产品化、服务定制化和生态共建三大核心方向构建，通过多维度收益模式实现商业化闭环，具体规划如下：

1. 拟定收入来源

硬件设备销售：面向短视频博主、智能农业从业者、摄影爱好者、监控安保部门等客户，提供标准化低功耗 1080p/4K 相机设备。当前项目 VIVIDIFY 摄像机整机成本控制在 400 左右元/台（自主可控的核心组件实现成本压缩），售价定为 1499 元/台，显著低于同类专业级 4K 拍摄设备，毛利率达 80%-300%。毛利润（Gross Profit）是指产品在一定时期内通过销售商品或提供服务所获得的收入，减去与生产或采购这些商品、服务直接相关的成本后的剩余金额。它是衡量产品核心业务盈利能力的重要指标。市场竞争力对比如下表格 5.1.1。

设备类型	成本范围	市场价格范围	备注
VIVIDIFY 摄像机	400 元/台	1499 元/台	低功耗，高分辨率
4K 单反相机	3000 元/台	8000 元/台	高功耗，高分辨率
专业级 4K 摄像机	6000 元/台	2 万元/台	高成本，高功耗，高分辨率

表 5.1.1 硬件设备市场竞争力对比表

软件服务订阅和技术支撑：提供视频超分边缘计算的云端服务订阅。定价模式为按需付费（Pay-as-you-go），基础功能免费（吸引生态用户），高阶画质（4K）升级服务（如亚米级精度）收费 20 元/月，购买我们产品的用户，可以免费体验一个月的视频超分服务。同时与华为、移动，联通，电信等企业合作，建立边缘计算站点，预计 3 年内覆盖全中国范围内所有设备。

2.成本结构：我们的生产成本占总成本的 45%，主要通过两个关键策略实现成本优化：一是采用 RISC-V 开源芯片架构，相比传统 X86 架构可降低 70% 的芯片成本，比起高性能芯片更加适合我们的铲平；二是通过 3D 打印一体化机身设计，将单台设备的成本控制在 100 元。研发投入占总成本的 30%，重点投向超分重构模型的性能优化。运营成本占比 20%，以北京为核心建立了覆盖全国的 24 小时应急响应服务体系。此外，我们还预留 5% 的生态建设预算，每年投入数万元支持 VIDIFY 摄像机社区，建立良好的社区环境，积极听取用户的意见并逐渐改进自身的产品。这种成本结构设计既保证了产品技术领先性，又实现了整体成本的有效控制。

3.利润模式：战略规划采用三阶段递进式发展路径：短期（1-3年）聚焦硬件销售业务，通过政府招标采购和运营商集中采购渠道实现正向现金流，确保 15%的净利润率目标；中期（3-5 年）推动业务结构转型，将软件订阅服务与数据增值业务营收占比提升至 50%，推动整体净利润率优化至 25%；长期（5 年以上）致力于构建"智能设备+核心算法+数据资产"的全栈式技术生态，通过参与传统相机企业合作、开展专利授权等高附加值业务，建立可持续的收益模式。这种阶梯式发展战略既保证了短期生存能力，又为长期技术壁垒构建预留了战略纵深。

4.合作伙伴：构建了三维立体化合作伙伴体系：在产业协同层面，联合北京邮电大学国家重点实验室、华为、通信运营商等头部机构成立"边缘计算云服务产业联盟"，实现订单资源共享与专利池交叉授权；技术生态层面，主导开源社区超分重构模型开发，通过全球开发者协同创新降低研发边际成本；市场拓展维度，积极响应"一带一路"倡议，积极拓展海外市场，计划到 2027 年实现国际市场收入占比 30%的战略目标。这种"产学研用"深度融合的伙伴网络，有效强化了项目在技术标准制定和市场渗透方面的双重竞争优势。

5.风险控制：为了有效控制风险，我们采取了一系列措施。在技术风险方面，我们预留了 20%的研发预算用于 8K 超分重构技术的前期研究，以确保在技术代际上保持领先地位。针对市场风险，我们与保险公司合作推出了“按需租赁”模式，例如为灾区制定单日万元的租金方案，以降低客户的采购门槛。在政策风险方面，我们积极参与了工信部《新一代人工智能发展规划》的制定工作，确保我们的产品和服务符合相关法规，同时塑造行业的话语权

5.2 运营模式

产品定位与目标市场

产品定位：VIVIDIFY 低功耗双模摄像机是一款创新型的视频拍摄设备，主要面向短视频创作者、智能农业从业者、野外监测研究人员以及智能安防从业者等。该产品通过独特的“低功耗黑白视频流+间歇式高分辨率关键帧”架构，能够在保证视频质量的同时，有效降低功耗，延长拍摄时间，满足用户在不同场景下的拍摄需求。

目标市场：短视频行业、智能农业、野外监测、智能安防等领域。

基于 VIVIDIFY 的技术特性与目标场景，设计「硬件+云服务+生态」三位一体的运营模式，实现商业闭环与社会价值双赢，我们的运营模式一共分为三个部分：

5.2.1 核心运营策略：分阶段渗透市场

阶段	目标	关键动作	营收重点
初创期（1-2年）	验证场景，建立行业标杆	<ul style="list-style-type: none">- B 端试点：与 MCN 机构/旅游局合作，免费提供设备换取案例- 开发者生态：开源基础 SDK，举办超分算法大赛- 云服务内测：邀请创作者试用云端处理服务	企业定制服务费 云存储增值费

成长期（3-4年）	规模化覆盖垂直行业	- 硬件订阅制：设备押金+按月付费（含基础云服务）- 政府采购：切入智慧农业/生态监测项目 - API 开放平台：向安防公司提供超分技术授权	订阅费用 技术授权费
成熟期（5年+）	构建低功耗影像生态	- 算力共享网络：用户出租设备闲置算力换积分- 数据增值服务：脱敏户外影像数据售予科研/文旅机构	数据交易佣金 广告分成

5.2.2 盈利模式设计：四维收入矩阵

1.硬件层（轻资产运营）

基础版：定价¥1499（双摄模组+4G 传输模块），成本控制在¥400

专业定制版：加装防水模块，防抖（¥2499+），与佳能，大疆等合作贴牌生产。

策略：通过订阅制降低购买门槛，设置 12 期无利润分期付款（首付¥499+月费¥89），提供按需付费。

2. 云服务层（核心利润点）

提供基础版，创作者版，企业版三个等级的云端超分服务。基础版功能包括 1080P 超分+100 小时存储，此服务版本购买 VIVIDIFY 低功耗超分相机便能够免费体验；创作者版可以享受如下功能：4K 超分+AI 色彩增强+500 小时，预计按照每年 ¥599 的收费模式进行订阅；企业版主要面向规模更大的创作者企业，订阅后可以享受到如下功能：多设备管理+API 接入+无限制存储。其收费标准预计 ¥2999/设备。具体信息如下表所示：

服务等级	功能权益	定价模型	目标用户
基础版	1080P超分+100小时云存储	设备免费赠送	个人创作者
创作者版	4K超分+AI调色+500小时存储	599元/年	职业博主/工作室
企业版	多设备管理+API接入+无限制存储+亚米级检测	2999元/设备/年	安防公司/农业科技企业

3.数据价值层

超分影像数据库：因为用户拍摄的视频数据会上传到云端服务器，我们可以筛选用户上传的视频数据，并将合适的视频数据保存下来。我们将筛选好的视频数据整理为一个数据集，将其作为一个专门用于基于关键帧的视频超分的数据集。因为有所有创作者共同构建，我们的数据集将具有丰富的**场景多样性**，覆盖野外、城市、人像、自然景观等 20+ 垂直场景。利用这样的数据集训练出来的超分模型将具有更高的稳定性和鲁棒性，向高技术企业、科研机构（中科院）提供订阅服务。

同时我们可以建立短视频**创作素材库**：用户可以上传高质量短视频素材，搭建一个可以供视频创作者互相交流互相学习的创作平台。用户上传的超分视频可进行自由交易，平台抽成 30%。我们会设立**创作者激**

励计划，优质素材将会被推荐至抖音/哔哩哔哩短视频内容池，曝光量转化收益分成。也会和企业合作，包括不限于采购特定场景素材，定制化视频内容等，提供创作者更多的合作机会。为保证创作者素材的创作版权，我们会联合太平洋保险推出盗版赔付服务，保障交易安全。素材库不仅可以解决创作者**高频优质素材需求**，还能通过用户上传数据优化超分算法。具体模式如下图所示：

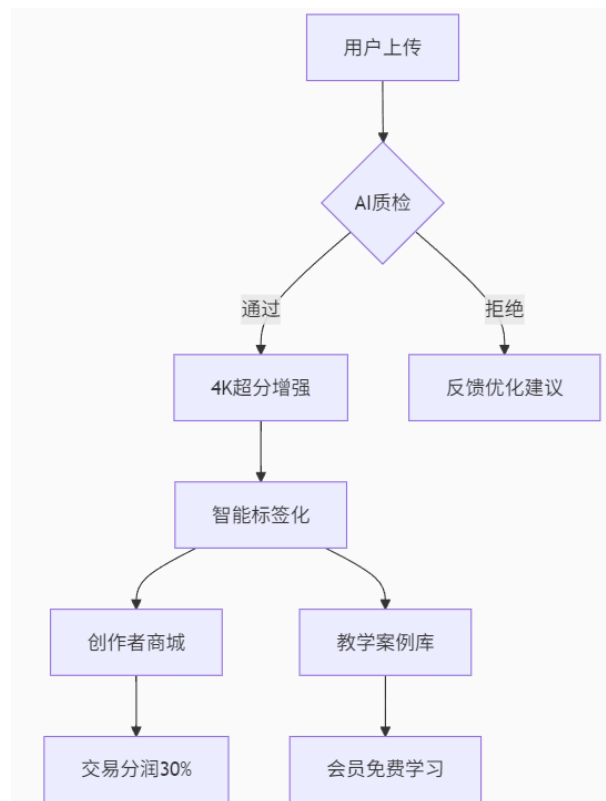


图 5.2-1 创作素材库运行模式

4.生态层

用户社区建设：建立用户社区平台，促进用户之间的交流与互动。鼓励用户分享拍摄经验、创作技巧、使用心得等，形成良好的社区氛围。如，联合专业摄影师与算法工程师，定期开展技术实训直播（如超分参数调优、动态场景拍摄技巧），提供**开源 SDK 工具包**（含光流对齐模块 API、自定义 LUT 调色模板），降低创作门槛。不局限于线上，同

时组织线下设备联测活动（如夜景拍摄挑战赛），供创作者一起交流讨论。开放硬件接口规范与模型训练框架，吸引第三方开发者贡献插件

用户激励与反馈：建立用户激励机制，对积极参与社区活动、创作优秀内容的用户给予奖励，如积分、荣誉称号、产品优惠等。定期收集用户反馈，根据用户需求和建议不断优化产品和服务。用户可以根据自身在社区的活动获得积分，如用户通过教程创作（千次播放=50 积分）、BUG 提交（有效反馈=100 积分）等行为获得积分，积分可以兑换更多的拍摄组件或者获得算力加速等服务。同时根据用户的贡献值设置社区等级，从低到高分别为“见习工程师→光影院士”头衔等等。以此来提高用户的活跃度，和完善社区营造良好的社区氛围。

5.2.3 渠道策略：三线并行

基于多元渠道协同拓展的战略布局，项目将通过**三位一体销售网络**全面渗透目标市场：**线上渠道**以官方品牌账号为核心阵地，同步搭建电商平台直营店铺，提供全链路产品展示与一站式购买服务；**线下渠道**在重点城市设立品牌体验店，配备动态场景演示区，用户可实地体验低功耗拍摄与超分画质增强效果；**政企合作渠道**深度响应国家乡村振兴战略，对接农业农村部高标准农田建设、生态监测等项目，为智慧农业场景定制太阳能供电版摄像机，推动数字技术在农业生产中的规模化落地。

六、教育维度

6.1 树立价值理念

本项目立足于国家“**绿色低碳，智能制造**”发展战略背景，深度对接《2030 年前碳达峰行动方案》与《新一代人工智能发展规划》实际需求，在技术研发的全过程中始终坚持“服务社会、赋能国家、低碳环保”

的理念引导，积极引导学生在科研活动中树立正确的价值观和社会责任感。我们聚焦低功耗视频拍摄领域，设计一种采用“低功耗黑白视频流+间歇式高分辨率关键帧”架构的高清摄像机系统，兼顾视频质量和系统节能。为短视频拍摄，智能农业监测，野外相机拍摄等多个领域，提供了创新性的，更高效的解决方案。

项目不仅让学生认识到前沿技术服务民生的战略意义，还引导大家在项目推进过程中不断反思技术的社会作用，从而完成从“技术导向”到“问题导向”“价值导向”的认知跃迁。在立项、中期到系统成型阶段，团队成员主动关注短视频行业发展、拍摄设备目前面临的困境与目前用户的拍摄需求等议题，展现出良好的家国情怀与工程伦理意识。

通过多轮研讨与实践，学生更加清晰地认识到：技术并非孤立存在，而是与社会、文化、伦理等多方面密切相关。这种理念贯穿整个研究过程，有力促进了学生综合素质与价值判断能力的提升。

6.2 推进专创融合

本项目紧密依托“视频拍摄”“人工智能”“超分重构”“嵌入式系统”等核心课程体系，围绕“专创融合”目标，将专业课程知识与创新能力培养深度融合。在项目各阶段，学生需综合运用移动通信人工智能、计算机原理应用、现代通信、计算机网络、机器学习等课程内容，并进一步拓展掌握深度学习、深度神经网络、超分重构模型、边缘计算、嵌入式实时视频采集、光流对其算法和变形卷积等前沿技术。

项目采用“任务驱动+模块迭代”的组织形式，各模块工作不仅要求学生具备扎实的专业基础，更需不断进行跨模块协同与方案优化，真正实现了“以创促学”“以研促专”的良性循环。例如，团队在搭建超分重构模型时，必须理解并实现基于关键帧的超分处理流程；而在构建硬件拍

摄部分时，则要熟悉嵌入式开发理解 STM32 运行原理，利用 STM 控制两个拍摄设备。这种对跨专业技能和系统性工程能力的综合要求，有效打破了课程内容之间的壁垒，推动了专业教育与创新实践的深度融合。

在这一过程中，学生不仅将课堂所学知识应用于系统架构与模块实现，还进一步锻炼了问题建模、用户视角分析、商业逻辑推演等综合创新素质，实现了从技术开发者到解决方案设计者的角色转变。项目打通了“课程知识—项目研发—应用竞赛”的转化通道，真正实现了理论知识的落地转化与应用拓展，显著提升了学生的问题解决能力、系统思维能力与跨领域创新能力。

6.3 助力人才培养

本项目团队由通信工程学院“英才班”学生组成，借助学校科研平台与大创计划平台，在高水平教师团队的指导下，形成了良好的“科研小组式协作—工程型能力驱动—多角色交替提升”的人才培养机制。在项目实施过程中，学生明确分工、互补合作，涵盖算法设计、系统搭建、平台调试、硬件配置、文档撰写、演示汇报等多种角色。整个过程高度还原真实工程环境，为学生走向社会、融入企业研发打下坚实基础。

项目历时一年，团队成员经历了从需求分析、架构设计、模块开发到系统集成与功能验证的完整工程闭环。每位成员均独立负责核心模块的功能实现，并定期开展组内 `code review` 与技术研讨，不仅技术能力大幅提升，也强化了沟通能力与项目管理意识。在训练过程中，多名同学已掌握 C++/Python 混合开发、超分重构模型构建和训练、嵌入式开发等复合技能，多人参与申报发明专利、科研立项或在校创新创业竞赛中获得佳绩。

更为重要的是，项目所倡导的“边学边研、边研边用”的理念，让学生在真实场景中建立了对科研与工程价值的深刻理解。在未来学术或就业路径选择中，学生更具备面向国家战略需求、解决实际问题的能力与信心，真正实现了“科研育人”“竞赛育人”的复合型创新人才培养目标。