|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | 研究目标 | 方法框架 | 实验数据集 | 主要贡献、创新点 | 性能评估 | 开源代码 | 未来展望 | 发表情况 |
| **TarDAL** | 红外和可见光图像融合，针对目标检测 | 双层优化公式，双对抗学习网络生成器和双鉴别器结构；  先做了两种图像的融合，然后输入到yolo v5网络中做目标检测 | M3FD基准数据集、TNO、Roadscene、Multispectral | **提出TarDAL网络结构和M3FD数据集**  1. \*\*双层优化公式\*\*：提出了一个新的双层优化模型，用于同时解决图像融合和目标检测的问题。这种公式允许在融合过程中考虑目标检测的需求，从而生成既适合视觉检查又适合计算机感知的融合图像。  2. \*\*目标感知的双对抗学习网络\*\*：设计了一个包含一个生成器和两个鉴别器的网络结构，其中生成器负责创建融合图像，而两个鉴别器（目标鉴别器和细节鉴别器）分别评估融合图像中的前景目标和背景细节。这种结构有助于保留红外图像中的目标结构信息和可见光图像中的纹理细节。  3. \*\*合作训练策略\*\*：开发了一种合作训练策略，通过反向传播算法同时优化融合网络和检测网络的参数。这种策略不仅提高了融合和检测的性能，而且提高了训练效率。  4. \*\*多场景多模态基准数据集（M3FD）\*\*：构建了一个包含多种环境、照明、季节和天气条件下的红外和可见光图像对的全面数据集。 | 目标检测的mAP | [https://github.com/dlut-dimt/TarDAL](https://github.com/dlut-dimt/TarDAL。) | 数据集扩展、算法优化、鲁棒性提升 | 2022年CVPR |
| CDDFuse | 多模态图像融合，包括红外-可见光和医学图像 | 双分支Transformer-CNN架构；Restormer、Lite Transformer、INN块；  两阶段训练，第1阶段训练Encoder, Decoder，重建自身图像。（自编码器）  第2阶段，生成融合图像。 | MSRS、RoadScene、TNO等 | **提出CDDFuse融合方法**  1. \*\*双分支Transformer-CNN架构\*\*：通过双分支结构分别处理全局和局部特征，以更有效地捕捉和融合多模态图像中的信息。  2. \*\*相关性驱动的特征分解\*\*：相关性驱动的损失函数，增强低频特征（模态共享信息）之间的相关性，同时减少高频特征（模态特定信息）之间的相关性，从而实现更有效的特征分解和融合。  3. \*\*全局和局部特征融合\*\*：CDDFuse通过特定的融合层将提取的全局和局部特征进行融合，以生成最终的融合图像。有助于在保留细节的同时，整合来自不同模态的互补信息。  4. \*\*广泛的应用场景\*\*：CDDFuse不仅适用于红外和可见光图像融合，还可用于医学图像融合等其他多模态图像融合任务，显示了其良好的泛化能力和应用范围。  5. \*\*统一的测量基准\*\*：CDDFuse提出了一个统一的测量基准，用于评估和比较不同融合方法在下游任务（如目标检测和语义分割）中的性能 | 8矩阵评估指标，评估融合后的图像的质量 | [https://github.com/Zhaozixiang1228/MMIF-CDDFuse](https://github.com/Zhaozixiang1228/MMIF-CDDFuse。) |  | 2023 CVPR |
|  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |