

A decorative graphic on the left side of the slide, consisting of four horizontal rows of light blue wavy lines. Each row contains approximately 10-12 peaks and valleys, creating a rhythmic, water-like pattern.

# 调研报告

刘伟, 2022.7.10

# 基本信息

论文题目	Target-aware Dual Adversarial Learning and a Multi-scenario Multi-Modality Benchmark to Fuse Infrared and Visible for Object Detection
研究方向	跨模态图像融合
会议/出版 年/分区	CVPR-Oral 2022   CCF-A   1 区
作者	Jinyuan Liu, Xin Fan, Zhangbo Huang, Guanyao Wu, Risheng Liu , Wei Zhong, Zhongxuan Luo
一作单位	大连理工大学-软件技术学院

# 拟解决的问题

- 问题:

以往的图像融合技术致力于提高融合质量 (强调寻找共同点)

而丢失了某些 **差异信息** (在后续目标检测任务中这些差异信息又十分重要)

比如红外图像的轮廓结构, 可见光图像的纹理信息

- 解决:

将图像融合和目标检测结合在一个 **双层优化公式** 中

设计了一个 **面向探测的生成-双对抗学习网络**, 用于 **红外图像与可见光图像融合**, 以"求同存异"



# 双层优化公式

用于解决两模态图像的 互利性融合

引用已有基础方法:

$$\min_{\omega_d} \mathcal{L}^d(\Psi(\mathbf{u}^*; \omega_d)), \quad (1)$$

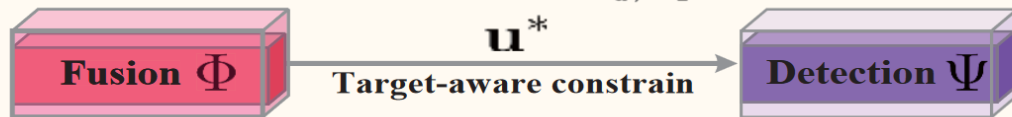
$$s.t. \mathbf{u}^* \in \arg \min_{\mathbf{u}} f(\mathbf{u}; \mathbf{x}, \mathbf{y}) + g_T(\mathbf{u}; \mathbf{x}) + g_D(\mathbf{u}; \mathbf{y}), \quad (2)$$

$\mathbf{x}$ 和 $\mathbf{y}$ 为源图像,  $\mathbf{u}$ 为融合图像,  $f(\cdot)$  是一个基于能量的保真度公式,  $g_t(\cdot)$  和 $g_d(\cdot)$  分别是在红外线和可见的两个可行性约束

由于上面公式不易求解, 改进: 引入可学习参数  $\omega_f$  来转化为单层

# 双层优化公式

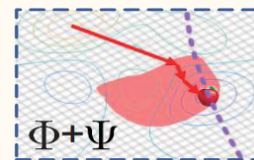
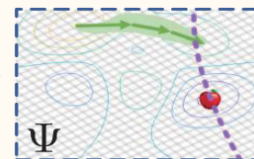
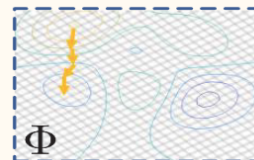
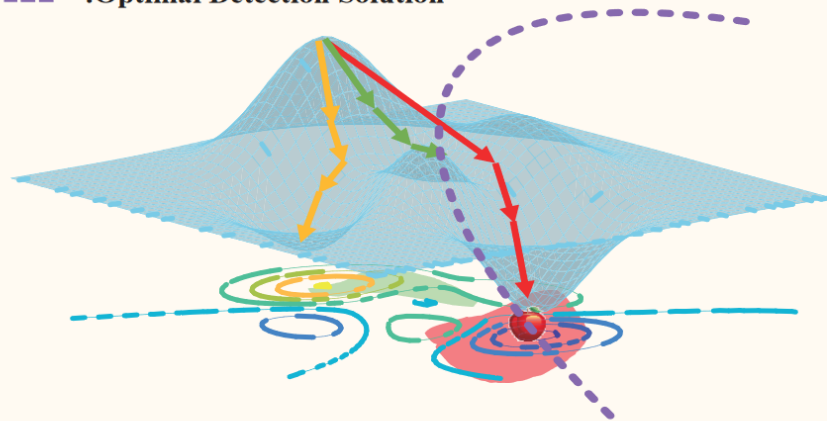
$$s.t. \mathbf{u}^* = \Phi(\mathbf{x}, \mathbf{y}; \omega_f) \quad \min_{\omega_d, \omega_f} \mathcal{L}^d(\Psi(\mathbf{u}^*; \omega_d))$$



:Optimal Fusion Solution



:Optimal Detection Solution

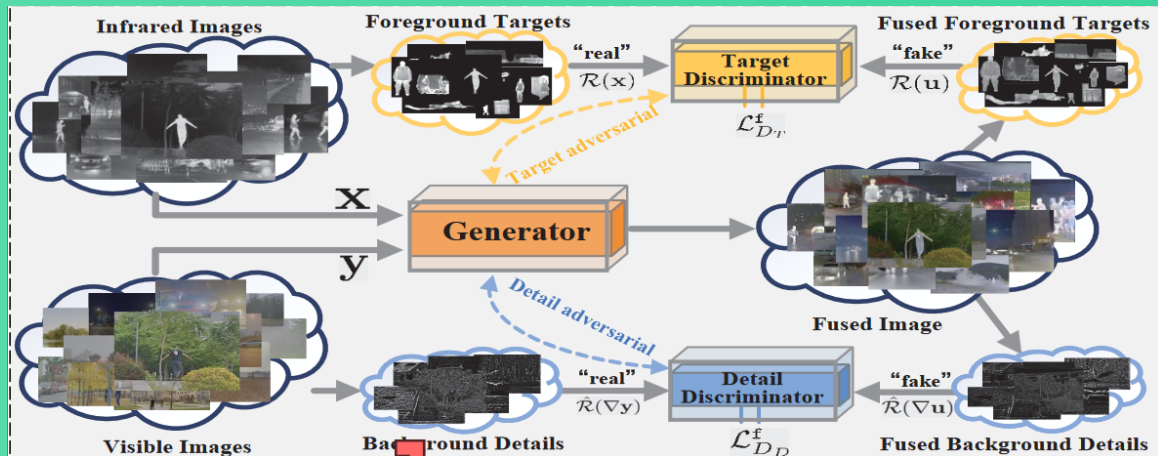


# 融合网络

检测网络  $\psi$  使用 YOLOv5 作为骨干

融合网络  $\phi$  为下面的 生成-双对抗网络，寻求共同点，同时从差异中学习

两个鉴别器  $D_T$  和  $D_D$  共享相同的网络结构，具有四个卷积层和一个全连接层

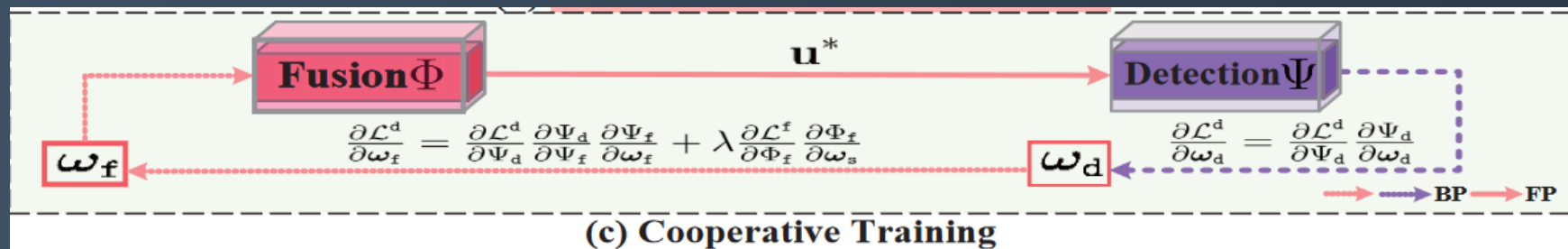


# 合作训练策略

双层优化自然地得出一个合作训练策略来获得最优网络参数  $\omega = (\omega_d, \omega_f)$

$$\min_{\omega_d, \omega_f} \mathcal{L}^d(\Psi(\mathbf{u}^*; \omega_d)) + \lambda \mathcal{L}^f(\Phi(\mathbf{x}, \mathbf{y}; \omega_f)), \quad (12)$$

$$s.t. \mathbf{u}^* = \Phi(\mathbf{x}, \mathbf{y}; \omega_f), \quad (13)$$







# 基本信息

论文题目	Unsupervised Misaligned Infrared and Visible Image Fusion via Cross-Modality Image Generation and Registration / 面向跨模态图像生成与配准的无监督未对齐红外与可见光图像融合
------	--

研究方向	跨模态图像融合与配准
------	------------

会议/出版年/分区	IJCAI 2022   CCF-A   1/2 区
-----------	----------------------------

作者	Di Wang, Jinyuan Liu, Xin Fan, Risheng Liu Luo
----	--

一作单位	大连理工大学-软件技术学院
------	---------------

- 问题:

现有大多数红外和可见光图像融合 (IVIF) 方法专门为手工制作的 **预配准图像** 设计

未对准的图像一旦存在轻微的偏差或形变就会导致融合图像中出现严重的重影现象

难以应对实际场景, 主要的障碍是 **跨模态的图像对齐**

- 动机:

迎合红外图像“重结构轻纹理”的固有特性, 提出了一种专门的跨模态生成-配准范式 (CGRP), 以减轻重影现象

# 论文贡献

- 跨模态生成-配准范式 (CGRP, 右图) 由三个子网络组成:

跨模态感知风格迁移网络 CPSTN

多级细化配准网络 MRRN

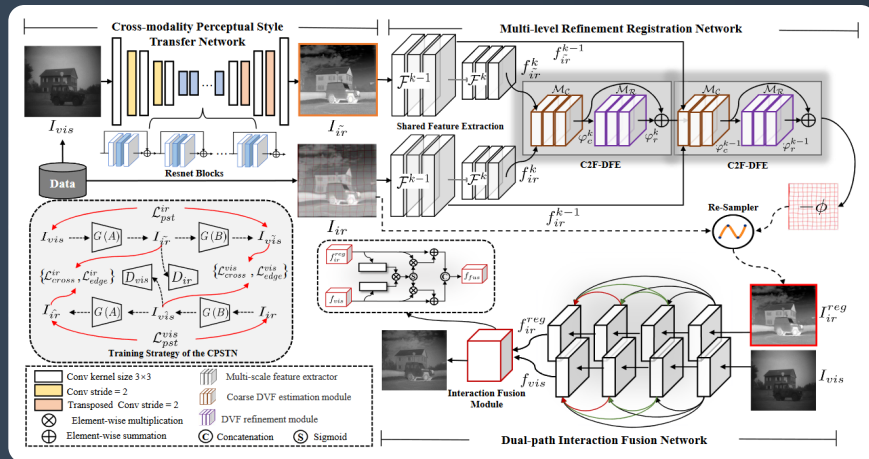
双路径交互融合模块 IFM

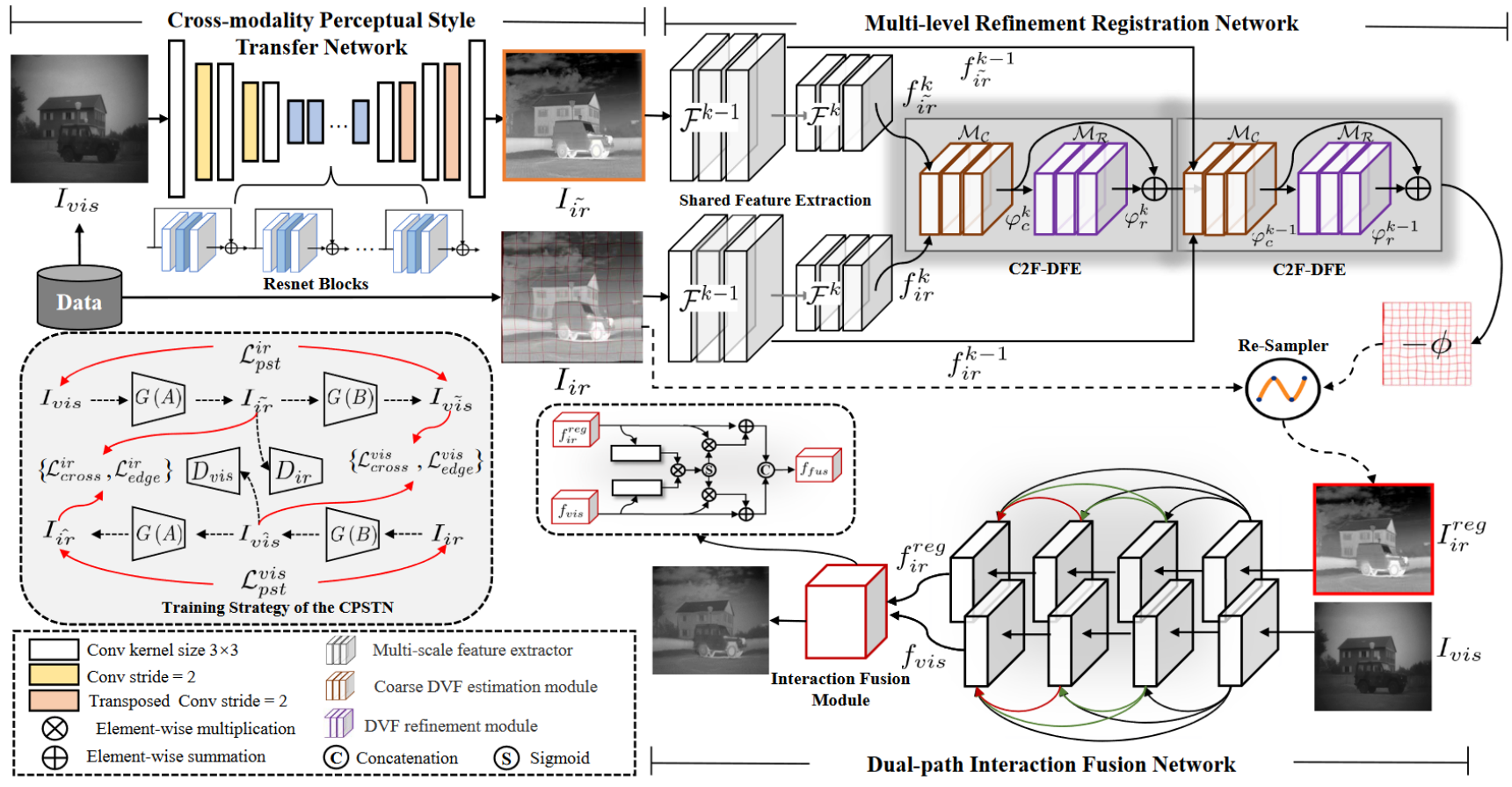
- 特点:

跨模态

无监督

未对齐的图像



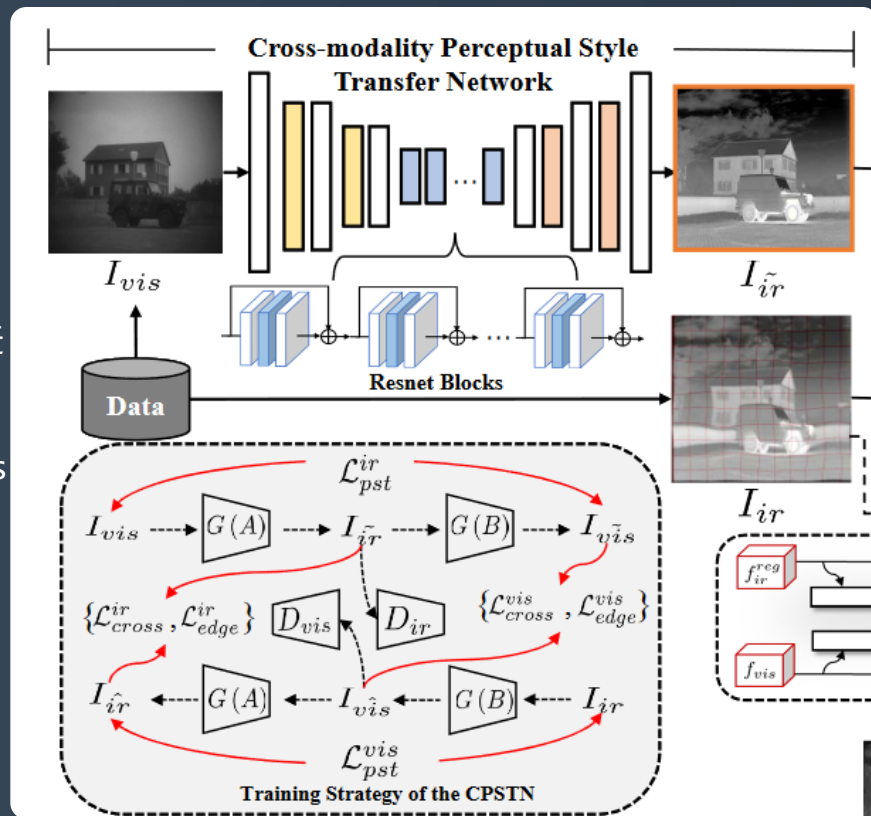


# 跨模态感知风格迁移网络 CPSTN

通过“可见图像到红外图像的转换”来降低跨模态差异

继承 CycleGAN 的基本学习方式, 不同的是使用由感知风格迁移约束控制的特定学习策略 (图下方)

此网络由两个生成器  $G(A) + G(B)$  和两个判别器  $D_{ir} + D_{vis}$  组成, 生成器网络如图上方为一个U形结构



# 多级细化配准网络 MRRN

- 此网络主要负责两红外图像的 **单模态配准**：

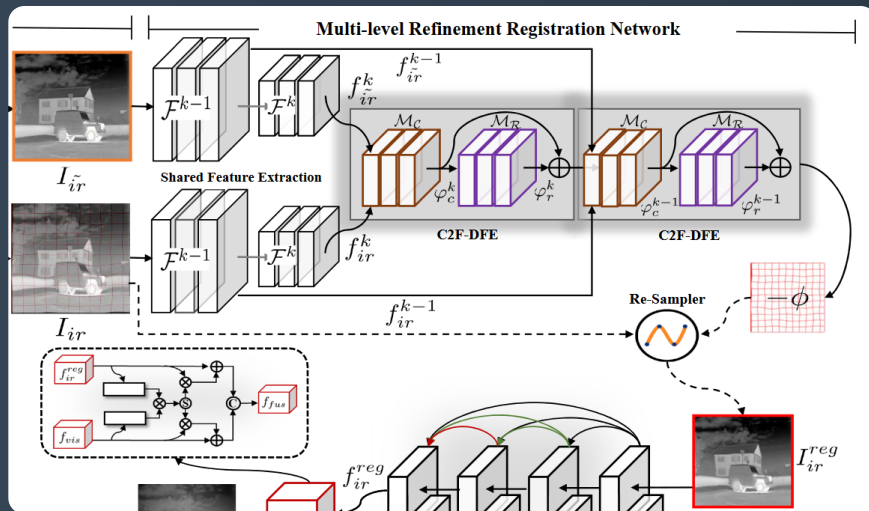
预测失真和伪红外图像之间的变形场 (DF) 并重建配准的红外图像

- 组成如下：

一个共享的多尺度特征提取器 (SM-FE)

两个由粗到细的变形场预测模块 (C2F-DFE)

一个重采样层



# 双路径交互融合网络

继承了残差密集网络, 并提出交互融合模块 (虚线框内)

