

工作总结与计划

一、工作总结（按时间划分）

日期	工作内容			完成情况及主要问题
	上午	下午	晚上	
星期一	读论文	读论文	读论文	了解论文要解决的问题，通过描述和插图认识了提出的网络架构及功能
星期二	处理杂事	读论文	读论文	简单学习了 Attention 机制，对网络中使用的部分方法的作用不太清楚
星期三	开会	读论文	读论文	阅读完论文对网络架构及其目的、作用的介绍和分析，对网络的设计有了更加清楚的认识
星期四	处理杂事	开会	休息	大致浏览了论文后续内容
星期五	读论文	写周报	开会	粗略阅读了实验结果和消融研究，对研究方向有了初印象
星期六	读论文	写周报	读论文	对第一篇论文进行了简单回顾，开始阅读第二篇论文

二、下一步计划（按任务划分）

编号	工作内容	目标	相关配合
1	完成第二篇论文的阅读	11 月 11 日前完成	无
2	准备毕设课题申报等工作	11 月 16 日前完成	待定
3			
4			
5			
主要风险	无		

三、个人分析与总结

内容提要	
1	进度方面：本周前半部分效率略高，按照预期计划进行，后半周动力不足，有待调整和改进
2	课题方面：通过阅读论文，对研究方向有了第一印象，并对一些深度学习和卷积的方法有了更多了解
3	其他思考：在阅读论文的过程中发现对实验结果部分的阅读缺乏耐心，可能需要再次认真阅读
4	

四、论文总结

论文标题	PMBANet:Progressive Multi-Branch Aggregation Network for Scene Depth Super-Resolution
作者及单位	Xinchen Ye, Baoli Sun, Zhihui Wang, Jingyu Yang, Rui Xu, Haojie Li, Baopu Li
论文出处	2020-TIP
创新点提炼	<div>1) 提出的并行的网络架构继承了集成学习的优势，可以从每个分支学习有效且多样化的特征。通过融合块（FB）自适应地选择和融合多个分支的特征中选择信息特征；</div> <div>2) 基于前馈/反馈和注意力机制设计了一种新颖的骨干网（RB），以增强高分辨率表示；通过 squeeze-and-expand 来提取高频特征（high-frequncy features），并进而 attention operation 来凸显提取出的高频特征；</div> <div>3) 通过密集连接和空洞卷积的重组构造单独的多尺度分支（MB），以更好地捕获多尺度信息；密集连接用于缓解梯度消失问题并利用从不同阶段提取的所有特征以获得更详细的多尺度信息；</div> <div>4) 在消融实验中，验证了颜色信息仅适合在较早的阶段引入以帮助深度重建，对 8×/ 16×情况有明显的帮助和性能的改善，但对 2×/ 4×却没有帮助更甚有害。</div>
个人想法	通过阅读论文，了解到颜色图片可以作为先验信息（但也有可能带来不需要的特征）帮助深度图像进行超分辨率重建，而在超分辨率重建的过程中，不同尺度的信息对于重建的精度有着至关重要的作用，因而多尺度和颜色信息可以帮助深度图像 SR。此外，论文中详细介绍了研究的驱动和设计思想，并通过多样的消融研究对网络架构中提出的多个模块（或创新点）进行性能研究，从而验证研究的合理性。

论文方法及结论：

（1） 论文提出的问题

- （尤其是在较大的放大倍数时）深度边界通常很难重建
- depth boundaries are generally hard to reconstruct particularly at large magnification factors
- 下采样退化严重破坏了场景中精细结构和微小物体上的深度区域
- depth regions on fine structures and tiny objects in the scene are destroyed seriously by downsampling degradation.

（2） 论文提出的解决方法

颜色引导的深度超分辨率被公式化为一个优化问题，它包括一个保真度项和一个先验项，以使不适定问题得到很好的约束。大致可以概括为以下优化函数：

color-guided depth SR is formulated as an optimization problem, which includes a fidelity term and a prior term to make the ill-posed problem well constrained. It can be roughly summarized into the following optimization function:

$$x^* = \arg \min_x \frac{1}{2} \|y - Kx\|^2 + \lambda \sum_l \omega_l * \rho_l(f_l \otimes x)$$

根据对现有算法及其不足等的分析（如手动生成的先验无法接近真实图像先验、现有的 CNN 降噪器并非专门针对深度超分辨率设计而使得颜色和多尺度信息的使用被忽略），本文追求更好的架构设计以进一步改善深度超分辨率重建。整个网络的设计方法都源自上述基于模型的优化方法，但同时使用深度网络对保真度项和先验项进行建模。

本文提出用于深度超分辨率重建的渐进式多分支聚合网络（progressive multi-branch aggregation network, PMBANet），由堆叠的多分支聚合（multi branch aggregation, MBA）块组成，以逐步恢复“降级”的深度图。

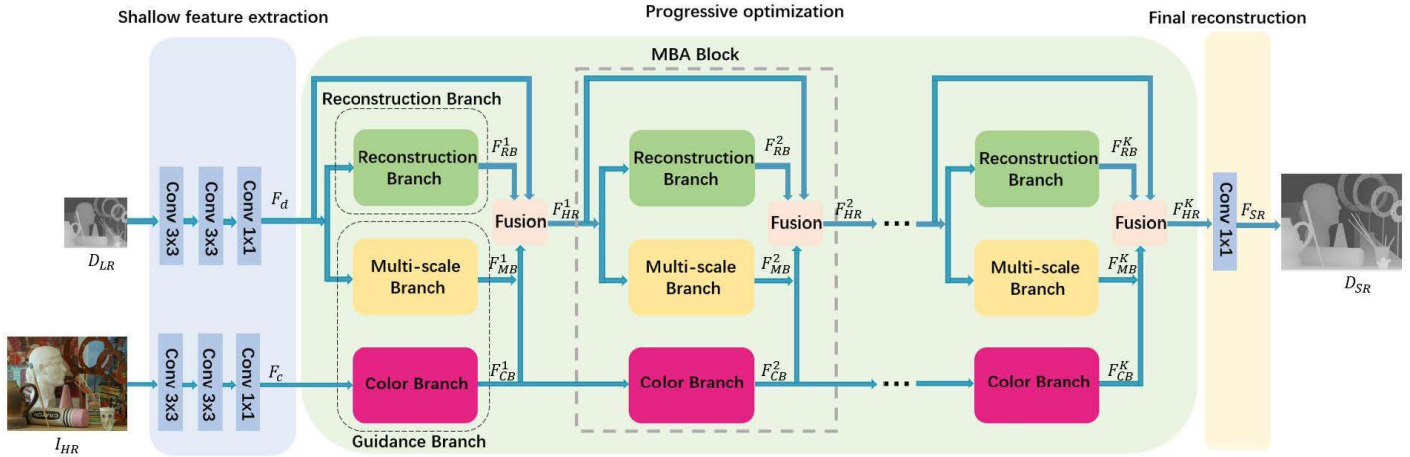


图 1 PMBANet 网络架构图

每个 MBA 块都有多个并行分支：

1) **重建分支（Reconstruciton Branch, RB）**，它是通过基于注意力的错误前馈/反馈模块（attention-based error feed-forward/back modules, AF）设计的，通过在前馈/反馈过程中引入注意机制以逐步突显深度边界的信息特征，从而迭代利用并补偿下采样误差以优化深度图；

2) **指导分支（Guidance Branch, GB）**，包括一个多尺度分支（Multi-scale Branch, MB）和一个颜色分支（Color Branch, CB），该分支为一个单独的子网以帮助 RB 恢复深度细节。

➤ MB 是要学习的多尺度表示，要密切关注不同尺度的对象。

➤ CB 则基于深度-颜色对之间的内部结构相关性，通过使用辅助颜色信息来对深度图进行正则化。

然后，引入融合块（Fusion Block, FB）从所有分支中自适应融合和选择特征。

整个 PMBANet 主要分为分为三个阶段：

- 1) **浅层特征提取（Shallow feature extraction）**：首先通过浅层卷积层分别从低分辨率的深度图（ D_{LR} ）和高分辨率的彩色图（ I_{HR} ）提取初始深度特征 F_d 和颜色特征 F_c 。 F_d 在下一阶段被输入到 RB 和 GB 中，而 F_c 仅被输入到 GB。
- 2) **逐步优化（Progressive optimization）**：逐步优化阶段由 K 个堆叠的 MBA 块组成。通过堆叠的 MBA 块，HR 特征空间中丢失的深度细节将逐渐得到恢复。
- 3) **最终重建（Final reconstruction）**：通过在最后一个 MBA 块的输出特征图上使用 1×1 卷积来超分辨重建深度图。

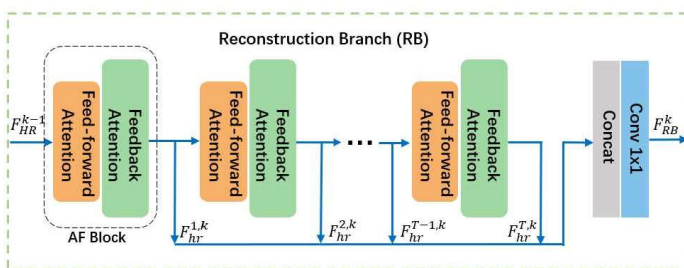


图 2 重建分支（Reconstruciton Branch, RB）

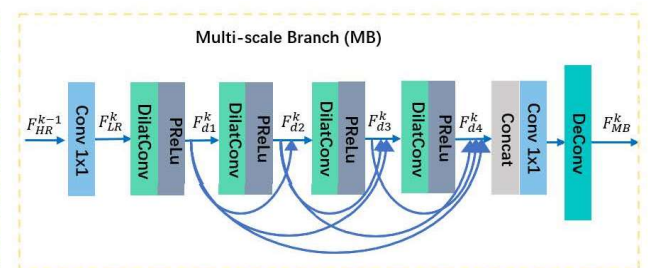


图 3 多尺度分支（Multi-scale Branch, MB）

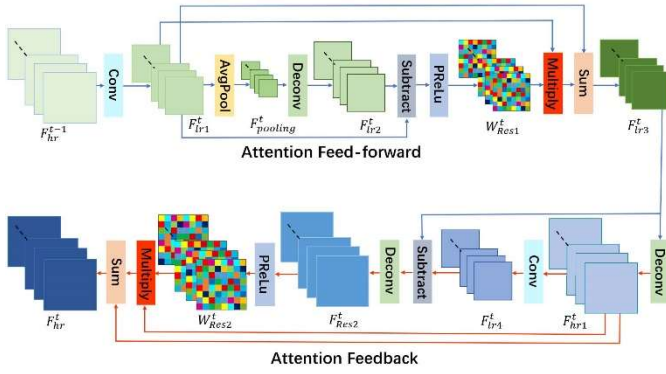


图4 基于注意力的错误前馈/反馈模块 (attention-based error feed-forward/back modules, AF)

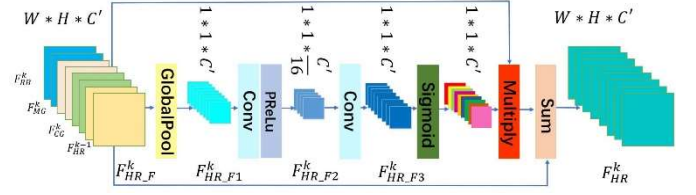


图5 融合块 (Fusion Block, FB)

在 RB 中, 受反馈机制和自注意力机制的启发, 提出了包括注意力前馈和反馈的 AF block (如图2所示), 注意力前馈突出显示 LR 域中的信息特征; 注意力反馈进一步增强了 HR 域中的有效特征, 通过注意前馈和反馈的交替, 逐渐减小重构误差, 从而更好地恢复深度细节。在 AF block 的两个阶段中, squeeze-and-expand 被用来提取高频特征 (high-frequency features), 而 attention 被用来凸显提取出的高频特征。

在 GB 中, 同时通过 MB 和 CB 有效地提取多尺度表示和颜色信息, 以帮助 RB 恢复深度细节。当解决下采样降级对多尺度对象的不同影响时, 有效的多尺度信息对于实现高精度深度超分辨率重建至关重要。如图3所示, MB 由四个扩张/空洞卷积及紧随其后的 PReLU, 1×1 卷积层和解卷积层 (DeConv) 组成。在 CB 中, 彩色图像可以用作先验信息以对 LR 深度图进行上采样。

FB 用于有效挖掘不同分支中特征通道之间的关系, 然后选择有用的特征通道进行深度 SR。通道注意力学习沿通道维度的图像特征权重分布, 并将学习到的权重应用于原始特征通道, 以专注于某些关键特征通道, 这适用于本文提出的 PMBANet 中 FB 需要自适应地聚合所有分支的不同特征通道的需求。

本文的主要贡献可总结如下:

- 1) 提出的并行的网络架构继承了集成学习 (ensemble learning) 的优势, 可以从每个分支学习有效且多样化的特征。通过融合块自适应地选择和融合多个分支的特征中选择信息特征;
- 2) 基于前馈/反馈和注意力机制设计了一种新颖的骨干网 (RB), 以增强高分辨率表示;
- 3) 通过密集连接和空洞卷积的重组构造单独的多尺度分支 (MB), 以更好地捕获多尺度信息;
- 4) 在消融实验中, 验证了颜色信息仅适合在较早的阶段引入以帮助深度重建, 在 $8 \times / 16 \times$ 情况下可以提供明显的帮助并显著改善性能, 但对容易恢复的 $2 \times / 4 \times$ 情况却没有帮助更甚至有反向作用。

(3) 实验结果结论

- 1) Middlebury Dataset (Noiseless Case): 为说明 PMBANet 的能力, 将其与其他最新的深度 SR 方法在不同的放大系数 ($4 \times$, $8 \times$ 和 $16 \times$) 下进行比较。与基于 CNN 的方法相比 (与基于 CNN 的方法相比, 基于传统过滤或优化的方法可获得相对较高的 MAD 和 PE 值, 即更差的性能), PMBANet 几乎获得了最佳的结果, 尤其是在 $8 \times$ 和 $16 \times$ 的情况下。

在场景中较大的物体上, 这些方法均取得了相似的结果, 但只有 PMBANet 在微小的物体上仍能准确地恢复深度边界。

- 2) Middlebury Dataset (Noisy Case): GSRPT 和 PMBANet 可以在消除噪音的同时保持最清晰的深度边界。

- 3) NYU Datasest: 在效率方面，PMBANet 在所有基于深度学习的方法中取得了最好的结果。
- 4) MPI dataset and Reak Data: 在大部分情况下取得了优于其他方法的结果。

(4) 存在的问题

在 CB 中，一些无用的纹理信息即使在最后的特征图中也无法被剔除，可能会导致导致纹理复制或深度图像中产生伪影（论文中已提到）。

其他问题暂时没有想到。