

# 基于希尔伯特空间曲线填充的太赫兹图像超分辨算法研究

杨墨轩<sup>1,2,3,4</sup>, 赵源萌<sup>1,2,3,4\*</sup>, 刘昊鑫<sup>1,2,3,4</sup>, 刘祎<sup>1,2,3,4</sup>, 吴悠<sup>1,2,3,4</sup>, 张存林<sup>1,2,3,4</sup>

1. 太赫兹光电子学教育部重点实验室, 北京 100048;
2. 太赫兹波谱与成像北京市重点实验室, 北京 100048;
3. 北京成像理论与技术高精尖创新中心, 北京 100048;
4. 首都师范大学物理系, 北京 100048)

**摘要:** 太赫兹成像技术受辐射源和探测器性能限制, 在细节分辨能力、成像速度和噪声抑制等方面仍有进一步改进的空间。本文提出一种基于空间曲线填充的太赫兹图像超分辨算法, 采用视觉自注意力(Vision Transformer, ViT)结构主干网络, 通过注意力机制进行太赫兹图像特征提取; 构建希尔伯特空间曲线, 根据特征图按曲线填充的方式进行图像重建, 并用轻量化的一维卷积处理重建图像特征, 对重建图进行逆变换恢复图像空间结构; 最终通过像素重组实现上采样, 得到物体轮廓和细节增强的输出图像。实验表明, 相较常规ViT结构, 本文方法图像峰值信噪比(PSNR)提高0.81 dB, 结构相似度(SSIM)提高0.0074, 有效抑制了噪声对图像纹理的影响, 获得了分辨能力显著提高的结果图像, 证明了太赫兹图像超分辨处理技术的可行性及其恢复图像细节、提高图像质量的能力。

**关键词:** 太赫兹成像; 超分辨重建; 注意力机制; 空间曲线

中图分类号: O436 文献标识码: A

## Terahertz imaging super-resolution algorithm based on Hilbert spatial curve filling

YANG Mo-Xuan<sup>1,2,3,4</sup>, ZHAO Yuan-Meng<sup>1,2,3,4\*</sup>, LIU Hao-Xin<sup>1,2,3,4</sup>, LIU Yi<sup>1,2,3,4</sup>,  
WU You<sup>1,2,3,4</sup>, ZHANG Cun-Lin<sup>1,2,3,4</sup>

1. Key Laboratory of Terahertz Optoelectronics, Ministry of Education, Beijing 100048, China;
2. Beijing Key Laboratory for Terahertz Spectroscopy and Imaging, Beijing 100048, China;
3. Beijing Advanced Innovation Center for Imaging Theory and Technology, Beijing 100048, China;
4. Department of Physics, Capital Normal University, Beijing 100048, China)

**Abstract:** The performance of radiation sources and detectors currently limits terahertz imaging technology, which still requires further improvement in terms of detail resolution, imaging speed, and noise suppression. This paper proposes a terahertz image super-resolution algorithm based on spatial curve filling. The ViT (Vision Transformer) structure backbone network is utilized to extract terahertz image features through an attention mechanism. A Hilbert spatial curve is constructed to reconstruct the image according to the feature map using the curve filling method. Lightweight one-dimensional convolution processing is used for reconstructing image features, while inverse transformation of reconstructed maps restores the image's spatial structure. Finally, pixel reorganization enables up sampling to obtain an output image with enhanced object contour and details. Experimental results show that compared with conventional ViT structures, this proposed method improves Peak Signal-to-Noise Ratio (PSNR) by 0.81 dB and Structural Similarity Index

收稿日期: 2023-10-14, 修回日期: 2024-04-21

Received date: 2023-10-14, Revised date: 2024-04-21

基金项目: 国家自然科学基金(61875140); 科技创新服务能力建设-高精尖创新中心-成像技术高精尖创新中心(19530012003); 校内专项--学位点建设与研究生教育质量提升(008-2355093)

Foundation items: Supported by the National Natural Science Foundation of China(61875140); Scientific and technological innovation service capacity building-Advanced Innovation Center - Advanced Innovation Center for Imaging Technology (19530012003); Campus Project -- Construction of Degree Sites and Improvement of Graduate Education Quality (008-2355093)

作者简介(Biography): 杨墨轩(1999—), 男, 北京人, 硕士研究生, 主要研究领域为太赫兹成像、光场成像、计算机视觉. E-mail: yang.ymx@qq.com

\* 通讯作者(Corresponding author): E-mail: zhao.yuanmeng@cnu.edu.cn

(SSIM) by 0.007 4, which effectively inhibits the noise influence on texture and significantly improves the resolution and image quality.

**Key words:** terahertz imaging, resolution enhancement, attention mechanism, space curve

## 引言

太赫兹波指波长在30至3 000  $\mu\text{m}$ 的电磁波,其对很多介电材料和非极性物质有良好的穿透性,而金属表面对太赫兹波有强反射性。太赫兹波光谱能量低,不会对生物组织造成电离损伤。太赫兹成像系统利用探测器采集太赫兹波的强度与相位信息进行成像,由于其良好的穿透性、安全性与隐蔽性,在国防军事、公共安全、航天遥感、无损检测等领域被广泛应用<sup>[1-4]</sup>。然而,受探测器性能、发射源功率、光学衍射极限、成像环境等限制,太赫兹图像存在分辨率低、纹理轮廓模糊、噪声干扰大等问题。为解决上述问题,近年来有较多基于图像超分辨率算法改善太赫兹图像质量、恢复图像细节、提高图像清晰度方面的研究报道,特别是基于深度学习的新方法不断涌现。

2019年,Kiarash提出一种基于点扩散函数的太赫兹成像分辨率增强算法<sup>[5]</sup>,该方法使用时域和频域中的滤波器用于滤除噪声、低频频谱和衍射失真。同时,对点扩散函数(point spread function, PSF)进行了建模,并通过卷积进一步提高了输出图像的质量。该方法有效改善了图像质量与细节信息,但对噪声敏感,导致边缘信息恢复不够理想。2021年,Wang等提出一种基于复合卷积神经网络的太赫兹图像超分辨率方法<sup>[6]</sup>,根据太赫兹波的性质,将卷积网络从实数域扩展到复数域,利用振幅和相位信息构建复合卷积网络(complex convolutional neural network, CCNN)重建高分辨太赫兹图像,该方法泛化能力优于传统卷积网络,但此架构参数通过MNIST数据集训练获得,依赖硬件提供多维数据,适用性有待提高;Lu等提出一种结合数学退化模型与卷积神经网络的太赫兹图像超分辨率方法<sup>[7]</sup>,数学退化模型考虑了影响太赫兹图像质量的三个可能因素:模糊核、噪声和下采样,训练卷积网络从退化模型中学习映射关系来重建高分辨率图像,实验表明该方法可显著提高图像质量,但结果图像对成像距离较敏感。2022年,Ruan等提出一种基于亚像素卷积神经网络的太赫兹图像超分辨率方法<sup>[8]</sup>,通过引入注意力机制让网络关注更多的信息特征,用亚像素卷积神经网络获得精度和视觉效果更佳

的高分辨太赫兹图像。其优势为在上采样过程中使用了全部信息,上采样图像每个像素点都具有唯一映射。缺点为图像中冗余众多,增加了计算成本,优化困难;Yang等提出一种采用结合注意力机制的太赫兹图像退化模型的超分辨率重建方法<sup>[9]</sup>,通过网络通道间依赖关系的自适应调整,使网络更侧重图像高频信息的恢复,从而使图像获得更好的边缘特征。

通过太赫兹图像超分辨率算法突破器件与系统性能对成像质量的限制,对于太赫兹图像处理的下游任务来说具有重要意义。本文提出一种基于希尔伯特曲线填充机制的太赫兹图像超分辨率算法,拟通过空间曲线机制改善常规视觉自注意力(vision transformer, ViT)结构的感受野,采用轻量化的卷积模块提高图像质量,增强重建图像细节与纹理特征,并突出图像边缘。下文将从原理、方法和实验等方面具体阐述。

## 1 超分辨率重建原理与方法

### 1.1 算法流程与网络结构

基于ViT的超分辨率网络架构具有恢复高质量图像的能力,但由于没有卷积的感受野机制,依靠先验位置信息较难恢复图像纹理信息,常致出现变形、边缘模糊等问题。图1为本文采用的基于希尔伯特空间曲线填充的太赫兹图像超分辨率网络模型图,其中图像编码部分使用经典ViT架构提取图像特征,解码部分利用像素重组(Pixshuffle)实现上采样重建图像,以期望恢复图像细节与纹理<sup>[10]</sup>。在此基础上,本文增加一组希尔伯特模块,对ViT提取的特征图进行希尔伯特变换后,通过轻量化的卷积块对一维特征信息进行处理,实现对图像细节信息的提取与增强,最后对图像做逆希尔伯特变换恢复图像二维空间信息。

### 1.2 注意力机制层

本文使用由Dosovitskiy等人提出的ViT架构作为图像编码器主干网络<sup>[11]</sup>,不同于卷积神经网络,注意力机制利用大量全连接层与残差结构设计,克服卷积神经网络无法无限增加网络深度与宽度的问题。太赫兹图像超分辨率算法针对主动式、被动式以及时域光谱扫描成像等多种太赫兹成像方式,实

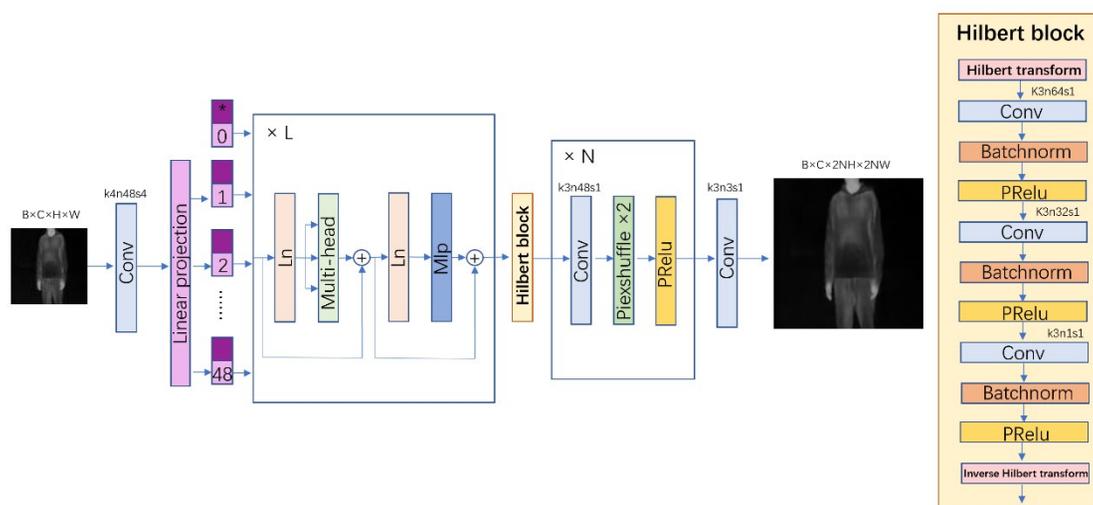


图1 基于希尔伯特空间曲线填充的太赫兹图像超分辨网络结构图

Fig. 1 Network structure diagram of terahertz image super-resolution based on curve filling in Hilbert space

际应用涵盖太赫兹人体安检、无损检测等多种场景,需以大量不同场景下的太赫兹图像构建数据集以供网络学习来提高泛化能力。文献[11]证明注意力机制的特征提取能力随着训练数据集的增大而提高,最终超过了卷积神经网络特征提取能力的上限。采用更多的神经元可有效突破卷积神经网络无法正确提取多样性数据特征的限制,有利于应对不同场景下的太赫兹图像超分辨任务。

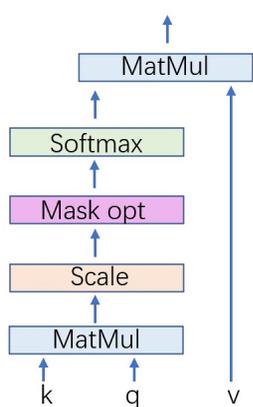


图2 注意力机制原理示意图

Fig. 2 Schematic diagram of the attention mechanism

注意力机制层不具有感受邻域信息的能力,针对这一问题,首先用大尺寸卷积核将图像分割为单个图像块并转换为向量形式进行拼接,然后为每一维向量增加初始的位置信息,最后将补充位置信息后的向量组输入注意力机制模块进行计算。注意力机制模块中包含线性归一化 (linear normaliza-

tion, LN)层、多层感知机 (multi layer perceptron, MLP)层与多头注意力机制 (multi-head attention)层,其中注意力机制层结构图如图2所示,通过三个独立的全连接层将补充位置信息后的向量 $k, q, v$ 赋予不同的初始值后表示为 $kW_i^k, qW_i^q, vW_i^v$ ,多头注意力机制为单层注意力机制的叠加,数学表达式为

$$Attention(k, q, v) = \text{soft max} \left( \frac{kq^T}{\sqrt{D_h}} \right) v \quad (1)$$

$$head_i = Attention(kW_i^k, qW_i^q, vW_i^v) \quad (2)$$

$$MultiHead(k, q, v) = \text{Concat}(head_1, head_2, \dots, head_h) \quad (3)$$

其中, $k$ 为查询向量, $q$ 为关注向量, $v$ 为特征值向量, $W_i^k$ 表示向量 $k$ 的第 $i$ 层权重;为防止通过激活函数的值过大导致梯度消失的情况出现,增加 $D_h$ 比例因子维持计算梯度。图2为单层注意力机制计算示意图,本文将多个单层注意力机制层输出结果叠加组成多层注意力机制层以获得更高维度的信息量。

### 1.3 希尔伯特空间曲线层

太赫兹图像超分辨率算法研究的一个关键之处在于如何在提高太赫兹图像分辨率的同时,保持对图像细节的结构感知 (Structure-Aware) 与平滑。在太赫兹图像中,结构感知与平滑需要考虑给定位置处图像的原始值与窗口内周围像素点的显著性差异,在保证图像结构信息正确的同时,通过优化相同属性像素点间的过渡方式得到平滑图像,减少噪声与畸变对图像质量的影响<sup>[12]</sup>。对于二维图像来说,需要借助卷积进行处理,在大尺寸与多通道

数的情况下,即使使用金字塔数据结构或者 Wino-grad<sup>[13]</sup>等卷积加速策略,仍会产生大量的计算成本。所以,本文提出一种利用一维空间曲线填充的方案处理太赫兹超分辨图像的边缘结构和细节。首先,通过空间曲线变换函数对图像做降维变换,表达式为

$$\{f(x,y,z)|x \in [0,h],y \in [0,w],z \in [0,c]\} \rightarrow \{F(x')|x' \in [0,w \times h \times c]\} \quad (4)$$

其中, $f$ 为太赫兹图像的映射, $F$ 为空间曲线变换函数, $h$ 为图像宽度, $w$ 为图像高度, $c$ 为图像通道数。空间曲线变换的关键思想是用一条不间断的曲线遍历图像中全部像素点,图像由此降维至一维。在图像降维过程中势必会有一些信息量损失,因此空间曲线变换函数的选择是保证图像邻域信息与结构感知的关键所在。常见的空间曲线有Z曲线、皮亚诺(Peano)曲线、希尔伯特(Hilbert)曲线等。希尔伯特曲线是一种豪斯多夫维为2的空间曲线,图3为希尔伯特曲线生成示意图。取边长尺寸为2的正方形来表征4个单位像素点,按图3(a)的排列方式进行连接得到一阶希尔伯特曲线;取边长尺寸为4的正方形并分为4个象限,其中第3象限顺时针旋转90度,第4象限逆时针旋转90度后按顺序连接构成如图3(b)所示的二阶希尔伯特曲线;以此类推取边长尺寸为 $2^n$ 的正方形来表征 $2^{2n}$ 个单位像素点,构成 $n$ 阶希尔伯特曲线。

本文使用邻域差异度来表征空间填充曲线在不同窗口范围内对图像信息的保持能力,表达式为:

$$R = \sum_{j=1}^{h \times w - k} \left\{ \frac{1}{k} \sum_{i=0}^k F(i+j) - F(i) \right\} \quad (5)$$

其中, $R$ 为曲线邻域差异度; $k$ 为邻域范围。本文对比了三种空间填充曲线在8邻域、24邻域、48邻域情况下的邻域差异度,对应尺寸分别为3、5、7的二维卷积核,结果表明希尔伯特空间填充曲线从二维向一维降维的过程中依旧保持了良好的结构与空

间信息,如表1所示。邻域差异度越小意味着降维后重新排列的像素点间保留了更多二维信息,这种额外保留的二维信息有效减少了邻域差异过大造成的混叠、割裂、边缘模糊等情况。

表1 不同曲线邻域差异度对比表

Table 1 Comparison of neighborhood retention rates of different curves

	8-Nbr	24-Nbr	48-Nbr
Z curve	2.132 3	3.193 3	4.409 0
Peano curve	2.176 4	3.462 9	4.274 3
Hilbert curve	1.815 9	2.836 5	3.979 0

因为希尔伯特空间曲线对降维图像结构与空间信息的良好保持,本文将其应用到网络优化过程中,构建如图1所示的希尔伯特块进行优化计算。第一步对图像做降维处理,降维虽然损失了图像部分空间信息,但减小了计算成本。当图像尺寸为 $w \times h \times c$ 时,需要进行 $w \times h \times c$ 次计算。第二步对图像做卷积、归一化、激活操作,一维的归一化与激活计算成本与二维相同,本文只计算卷积操作的计算成本。利用卷积核为 $ks$ 、步长为1的卷积进行计算,在二维图像尺寸不变的情况下需要进行 $w \times h \times c \times ks^2$ 次计算,而降维后仅需 $w \times h \times c \times ks$ 次计算。最后通过逆希尔伯特变换对图像进行升维处理,恢复图像空间信息,时间复杂度与降维一致,需要 $w \times h \times c$ 次计算。综上所述,基于希尔伯特空间曲线填充在单次卷积的情况下减少了 $w \times h \times c \times (ks^2 - ks - 2)$ 次计算,计算成本极大缩减,有利于模型优化以及太赫兹系统的实时应用。

#### 1.4 图像重构与损失函数

传统太赫兹超分辨算法是建立从低分辨率图像到高分辨率图像的一对多映射,这其中可能产生多个解,而确定最优映射成为能否生成高质量图像的关键。在太赫兹图像超分辨过程中,一部分高频数据是冗余信息,而图像中的低频分量包含了更多目标位置信息,可通过低频分量实现对图像的超分

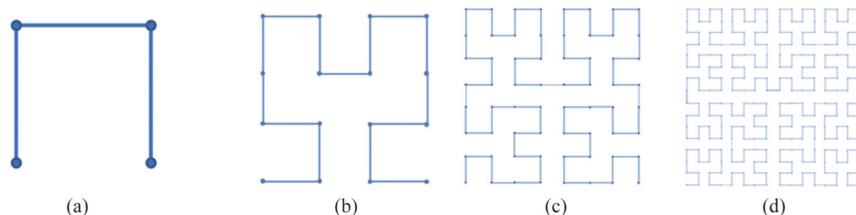


图3 希尔伯特空间曲线生成示意图

Fig. 3 Hilbert space curve generation diagram

辨重建<sup>[14]</sup>。本文采用像素重组(PixelShuffle)实现图像上采样,流程图如图4所示:

$$r^{SR} = \lambda_{spc1} \times M_c(I^{LR}) + \lambda_{spc2} \times HFP(M_c(I^{LR})), \quad (6)$$

其中, $M_c$ 为图像特征提取器, $M_c(I^{LR})$ 为特征图,HFP为图像高通滤波器, $\lambda_{spc}$ 为高频分量权重参数, $I^{LR}$ 为低分辨率图像(通过高分辨图像 $I^{HR}$ 下采样得到), $I^{SR}$ 为重建图像。在全部特征图中包含图像的低频特征与高频特征,其中高频特征决定了重建图像细节部分能否被清晰重建,边缘是否出现混叠。根据文献[14]的论述,像素重组技术通过建立各特征通道与重建图像间的映射关系,最大程度地保留了全部特征图的信息量,有效避免了转置卷积提高分辨率带来的畸变和边缘模糊等问题。为更大程度上增强图像锐度,本文在重建图像过程中增加了一组高通滤波器,期望通过残差连接高频特征图与原始特征图,以恢复图像细节与纹理。

为了使重建图像细节尽可能地接近真实图像,在训练损失上本文用均方误差(mean squared error, MSE)来保证图像在优化过程中可以更好地建立从低分辨率图像到高分辨率图像的映射,数学表达式为:

$$L_{MSE} = \frac{1}{c^2hw} \sum_{x=1}^{cw} \sum_{y=1}^{ch} (I_{x,y}^{HR} - M(I^{LR})_{x,y})^2, \quad (7)$$

其中, $L_{MSE}$ 为均方误差损失值, $M$ 为太赫兹超分辨模型, $h$ 为图像宽度, $w$ 为图像高度, $c$ 为图像通道数, $I^{LR}$ 为低分辨率图像, $I^{HR}$ 为高分辨图像。仅使用均方误差合成的重建图像在图像质量客观评价指标上表现优异,但其图像纹理并不符合人眼视觉习惯,轮廓纹理过于锐化<sup>[15]</sup>。为了获得更加接近人眼视觉的重建图像,本文将均方误差与几何视觉(visual geometry group, VGG)网络中间层特征图的绝对值损失组成复合损失函数,数学表达式为:

$$l_{VGG} = \sum_{i=1}^3 \lambda_i \|\phi_i(I^{HR}) - \phi_i(M(I^{LR}))\|_1, \quad (8)$$

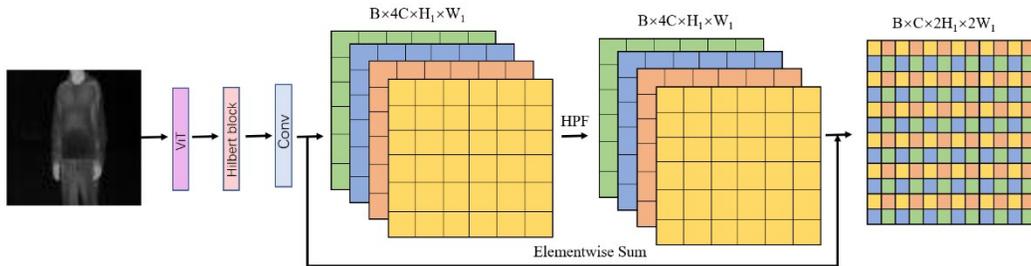


图4 像素重组流程示意图

Fig. 4 Schematic diagram of pixel reorganization process

$$l_{SISR} = \lambda_{MSE} l_{MSE} + \lambda_{VGG} l_{VGG}, \quad (9)$$

其中, $\phi_i$ 为第*i*层卷积块, $l_{VGG}$ 为VGG损失值, $\lambda_{MSE}$ 为均方误差损失权重, $\lambda_{VGG}$ 为VGG特征图损失权重, $l_{SISR}$ 为复合损失值,用此复合损失函数进行重建。

## 2 太赫兹图像超分辨重建

### 2.1 实验环境与训练参数

本文实验使用RTX4090型GPU,采用可见光图像数据与太赫兹图像数据的迁移学习策略。其中,可见光图像100 000张,来自SA-1B公开数据集<sup>[16]</sup>。太赫兹图像10 000张,实验数据采集设备如图5(a)所示,使用INO公司的MICROXCAM-384i-THz型太赫兹焦平面阵列相机采集原始太赫兹图像,采集方式如图5(b)所示,采集目标人物保持运动,对采集视频做逐帧处理。测试数据全部为太赫兹图像,共3 000张。在训练过程中使用随机剪裁的策略进行数据增强,输入高分辨图像尺寸为512\*512( $I^{HR}$ ),通过两次最大池化层降低分辨率至128\*128并作为低分辨率图像( $I^{LR}$ )输入模型。优化器使用Adam,学习率从 $10^{-3}$ 线性递减至 $10^{-5}$ ,基于ViT的编码器训练轮数100轮,希尔伯特层与重构层训练轮数1 000轮,批量处理尺寸32张。损失函数中使用VGG-19作为特征提取器,同时 $\lambda_{MSE}=0.7$ , $\lambda_{VGG}=0.3$ ,通过提取前三层VGG特征优化模型纹理细节。

### 2.2 实验结果

太赫兹图像超分辨研究主要受太赫兹图像数据采集困难、数据量小与数据多样性不足的影响。太赫兹成像得到的图像分辨率受波段特征和成像器件性能影响较大。基于时域光谱扫描的太赫兹成像系统成像精度高,但成像成本过高且无法实现实时成像。而焦平面阵列太赫兹相机采集帧率高,在实际成像应用中具有可实时成像的优势,因此本文选择用太赫兹相机采集原始图像。针对太赫兹

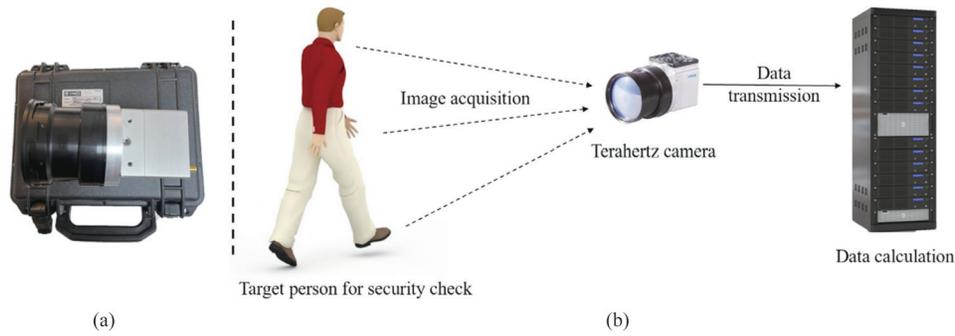


图5 实验环境示意图,(a)实验设备;(b)图像采集方式

Fig. 5 Schematic diagram of the experimental environment, (a) experimental equipment; (b) image acquisition

图像数据种类多样性不足的问题,本文首先基于100000张来自SA-1B数据集的可见光图像训练ViT的编码器架构,以保证其可以更加准确地提取太赫兹图像的特征。在编码器训练完成后冻结编码器权重参数,利用充分学习可见光图像的编码器提取太赫兹图像特征值并重建超分辨率图像。为验证本文提出的希尔伯特空间曲线填充机制对图像特征编码的有效性,对比加入希尔伯特空间曲线层前后的模型在重建图像中的MSE损失,结果如图6所示。可见,加入希尔伯特空间曲线层对特征进行编码,有效提高了模型的泛化能力,通过多个一维卷积编码保证了模型使预训练特征能够更有效捕捉太赫兹图像中的关键信息。在训练集中测试有无希尔伯特空间曲线层的结果差异不大,但在测试集中加入了希尔伯特空间曲线层的模型具有更好的对抗过拟合的能力。本文认为对编码器获得特征的二次编码设计给其他太赫兹图像下游任务提供了一种可行的设计思路。

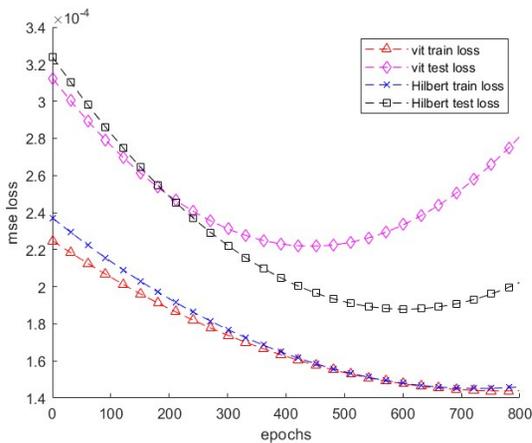


图6 希尔伯特空间曲线层训练与测试损失对比图

Fig. 6 Hilbert space curve layer training and test loss comparison diagram

在加入希尔伯特模块后模型额外引入了3层具有可学习参数的卷积层,虽然本文使用图像降维方法,将 $3 \times 3$ 卷积层可学习参数减少 $1/3$ , $5 \times 5$ 卷积层减少 $1/5$ , $7 \times 7$ 卷积层减少 $1/7$ ,减少了 $w \times h \times c \times (ks^2 - ks - 2)$ 次计算,但更多的可学习参数可能会间接影响模型性能。为验证希尔伯特空间曲线填充机制的有效性,本文对比ViT模型、卷积核尺寸分别为3、5、7时带有希尔伯特变换的ViT模型,以及卷积核尺寸分别为3、5、7时无希尔伯特变换的ViT模型的均方误差(mean squared error, MSE)、峰值信噪比(peak signal-to-noise ratio, PSNR)与结构相似性(structural similarity, SSIM)三种客观图像质量评价指标,结果如图7至图9所示。随着模型的不训练与优化不断降低,至700轮附近时均方误差下降到极小值,随着训练的继续均方误差逐渐增加,模型出现了过拟合的现象,泛化能力降低。训练至1000轮时模型出现退化现象,损失上升趋势增加,表明在此数据量下700轮时为模型最优参数,达到了最佳性能。另外,具有希尔伯特层的模型呈现出更优的超分辨性能,相较小尺寸卷积层,更大尺寸的卷积核对空间中邻域信息的感知能力更强,在客观评价指标上表现更好。没有加入希尔伯特变换机制的模型虽然增加了额外的可学习参数,但在重建图像能力上并没有超越原始ViT模型,部分模型出现了退化现象,实验结果如表2所示。可见希尔伯特变换是空间曲线填充层能够正确学习图像特征的必要条件,在没有加入希尔伯特变换的情况下,即便加入额外的可学习参数,模型也无法学习到更多的特征。为应对由于降维导致的空间信息丢失,可用更大尺寸的卷积帮助模型提取空间信息,替代二维卷积的感受野特性。

为验证本文提出的希尔伯特空间曲线填充算法用于太赫兹图像重建的可行性和有效性,对比

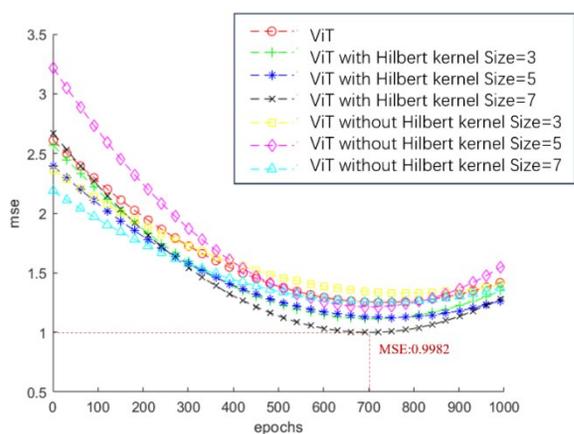


图7 不同模型均方误差对比图

Fig. 7 Comparison of mean square error of different models

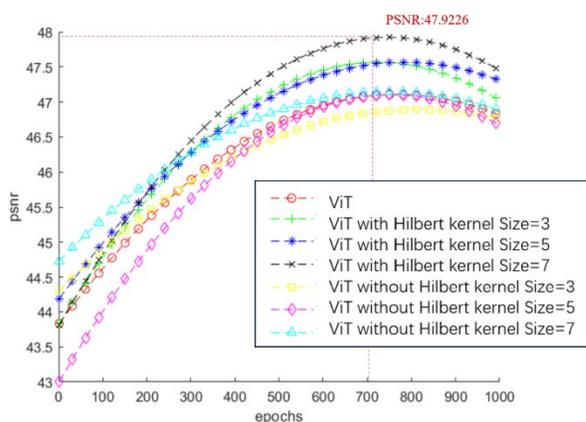


图8 不同模型峰值信噪比对比图

Fig. 8 Comparison of peak signal-to-noise ratio of different models

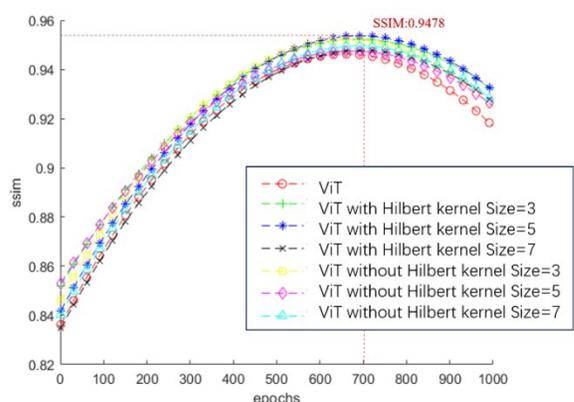


图9 不同模型结构相似度对比图

Fig. 9 Structure similarity comparison diagram of different models

ViT模型、无希尔伯特变换、有希尔伯特变换三种模型在不同尺寸卷积核下重建图像的能力,并分别提

取图像细节部分进行对比实验。实验结果如图10所示,图中的(a)原始图像,是用INO公司的MICROXCAM-384i-THz型太赫兹焦平面阵列相机采集的图像进行向下池化得到的;(b)~(h)是对(a)分别进行不同尺寸希尔伯特卷积核的图像超分辨重建结果;(i)是用INO公司的MICROXCAM-384i-THz型太赫兹焦平面阵列相机采集的未经向下池化的图像。在重建超分辨图像过程中,经典ViT模型与无希尔伯特变换的模型都出现了不同程度的过拟合,真实高分辨图像中存在的合理噪点被引入,同时对图像噪声十分敏感,在重建图像中无法有效区分有效信息与噪声信息,导致大量噪点被引入。而加入了本文提出的希尔伯特曲线填充层的模型则可以有效区分噪声与前景信息,能够合成更加平滑的太赫兹重建图像。在手部等细节部分的图像重建上,大尺寸卷积核表现出了更优秀的性能,获得了边缘清晰、更逼真的手部细节。但当没有希尔伯特变换层时,模型由于失去了空间信息无法通过手部低分辨率的低频信息预测高分辨图像中的高频信息,重建图像轮廓不够分明。实验结果表明,本文提出的基于希尔伯特空间曲线填充方法在太赫兹图像中的噪声敏感区域与边缘细节区域都具有较强的可靠性。

在对人成像问题上,不同波段成像方式间有一定的差异,如被动式太赫兹与红外波段成像通过采集人体自身辐射的电磁波进行成像,而可见光或X射线波段成像依靠外部辐射源直接或间接照射成像目标进行成像,成像方式影响所得图像的质量。在被动式太赫兹成像中,由于辐照度不足、探测器性能有限以及实际应用场景环境复杂等原因,目前获得的图像信噪比较低、边缘模糊、细节纹理信息有较多丢失,而可见光成像与X射线成像通常可有效捕捉更多细节纹理,获取更高质量的图像。图像波段不同、增强任务不同,适用的处理方法也不相同。为比较希尔伯特变换对太赫兹、红外、可见光与X射线图像超分辨处理效果的影响进行对比实验,实验结果如图11所示。在实验中太赫兹图像原始数据使用太赫兹安检图像数据集,红外与可见光图像来自LLVIP数据集<sup>[17]</sup>,X光图像来自Chest-Xray8数据集<sup>[18]</sup>;编码器使用基于ImageNet1K的预训练权重,训练与测试数据统一使用与Chest-Xray8数据数量相同的1125张;微调方式采用两种方案——全参数微调,以及更新解码器参数、冻结编码

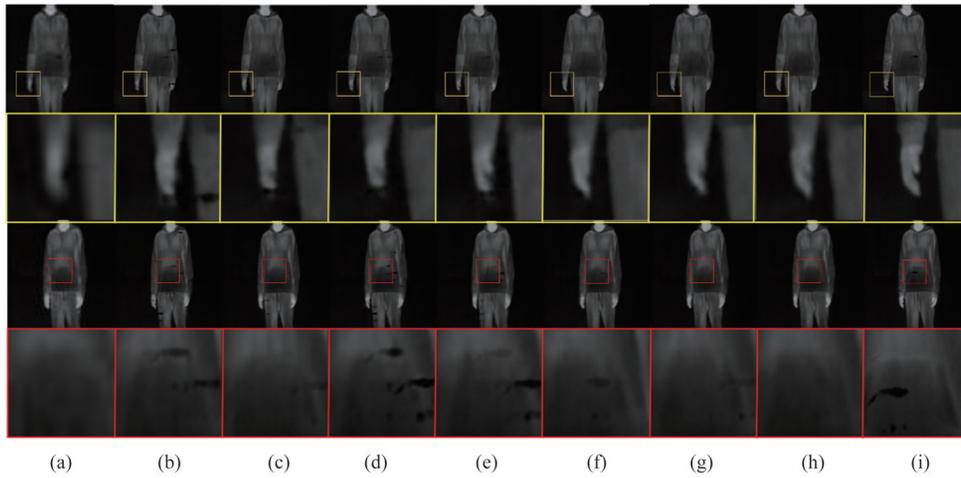


图 10 不同模型太赫兹图像超分辨重构对比图,(a)原始低分辨率图像;(b)ViT;(c) ViT无希尔伯特变换,卷积核=3;(d) ViT无希尔伯特变换,卷积核=5;(e) ViT无希尔伯特变换,卷积核=7;(f) ViT有希尔伯特变换,卷积核=3;(g) ViT有希尔伯特变换,卷积核=5;(h) ViT有希尔伯特变换,卷积核=7;(i)真实高分辨图像

Fig. 10 Superresolution reconstruction comparison of terahertz images of different models, (a) original low-resolution image; (b) ViT; (c) ViT without Hilbert kernel size=3; (d) ViT without Hilbert kernel size=5; (e) ViT without Hilbert kernel size=7; (f) ViT with Hilbert kernel size=3; (g) ViT with Hilbert kernel size=5; (h) ViT with Hilbert kernel size=7; (i) real high-resolution image

表 2 重构图像质量客观评价指标

Table 2 Super-resolution image quality evaluation index

Different algorithm	MSE	PSNR/dB	SSIM
ViT	1.251	47.108 7	0.946 3
ViT without Hilbert kernel size=3	1.331 4	46.892 9	0.952 1
ViT without Hilbert kernel size=5	1.213	47.102 6	0.947 5
ViT without Hilbert kernel size=7	1.254 5	47.149 8	0.949 6
ViT with Hilbert kernel size=3	1.113 5	47.564 2	0.952 4
ViT with Hilbert kernel size=5	1.123 5	47.566 2	0.953 7
ViT with Hilbert kernel size=7	0.998 2	47.922 6	0.947 8

器参数。实验结果表明,在加入希尔伯特变换层后,全局参数微调与仅微调解码器参数两种微调策略都在太赫兹图像与红外图像超分辨处理中获得了明显效果,客观评价指标 PSNR 也有较大提升。可见光图像分辨率增强处理在加入希尔伯特变换层后出现了一定程度的算法性能减弱,X光图像在多次实验验证中使用全局微调可带来部分提高,使用局部参数微调则出现了模型退化。由上述实验结果可知,希尔伯特变换能保持良好的降维图像结构与空间信息,同时还能表征出完全不同于可见光图像的目标物空间位置信息。实验过程中即使使用全局微调方案也难以改变 ViT 架构中位于网络上层的位置信息编码矩阵,本文算法的希尔伯特变换层引入全新的空间信息,弥补了 ViT 架构对纹理信

息量较小的太赫兹图像与红外图像编码能力不足的短板;而在冻结编码器仅更新解码器的方案下,也可通过在网络下层构建希尔伯特空间信息提高图像编解码能力,增强算法超分辨重构的能力。另外,本文算法受数据量与计算成本的限制小,还解决了因图像空间信息特征差异大无法使用可见光预训练权重的问题。

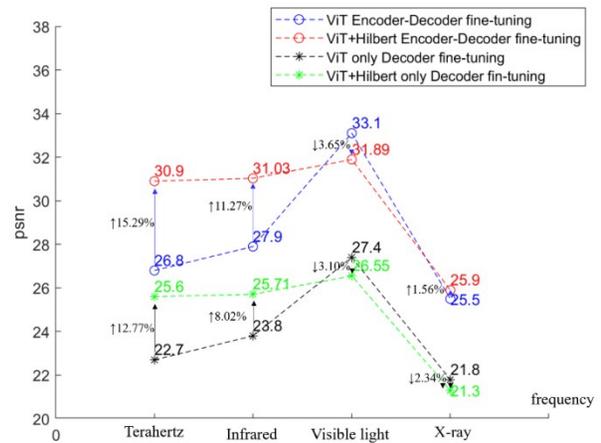


图 11 希尔伯特变换对不同频率成像方式影响对比图

Fig. 11 Comparison diagram of Hilbert transform effect on different frequency imaging methods

如图 12 所示,本文分别对比了邻域插值法(Nearest)、双三次插值法(Bicubic)、卷积超分辨网络

(super resolution convolutional network, SRCNN)<sup>[19]</sup>、残差超分辨网络(super resolution residual convolutional network, SRResNet)<sup>[20]</sup>、生成式超分辨网络(super-resolution using a generative adversarial network, SRGAN)<sup>[15]</sup>、亚像素卷积神经网络(efficient sub-pixel convolutional neural network, ESPCNN)<sup>[8]</sup>、ViT<sup>[11]</sup>、退化模型(mathematical degradation model, MDM)<sup>[7]</sup>的太赫兹图像超分辨重构能力。实验结果对比如表3所示,其中Nearest方法与Bicubic方法由于无法建立一对一的映射关系,无法有效恢复真实的图像信息。SRCNN与SRResNet由于其仅使用MSE损失,仅关注了图像像素级的损失,导致边缘模糊,纹理细节不清晰。SRGAN由于使用VGG中间层特征图的绝对值与MSE组成复合损失,在重构图像结构上有较大提升,重构图像符合人眼直觉,SSIM指标表现优异。但由于两种损失复合使用,导致图像噪点更多,图像不够平滑。ESPCNN引入了注意力机制应用于太赫兹图像编码,但由于其网络设计复杂,网络深度受限制,影响图像编码能力。

MDM模型将退化模型应用于太赫兹超分辨,其原理更符合传感器采集信号并合成图像,在图像重建过程中表现出优秀的重建能力。退化模型全局建模能力优秀,但由于其缺乏对图像细节部分的局部建模能力,导致图像细节部分出现畸变的现象。本文提出的基于希尔伯特空间曲线填充方法,在不需要对ViT编码器增加复杂网络结构的前提下,通过希尔伯特曲线特有的空间位置特征,在保持图像结构信息的前提下,克服了噪声对重构图像的影响,证实了本文方法在重构更好质量太赫兹图像方面的优势。

### 3 结论

针对太赫兹图像分辨率和信噪比较低影响图像质量的问题,本文提出了一种基于希尔伯特空间曲线填充的太赫兹图像超分辨方法。鉴于注意力机制虽具有强大的特征表示能力,但因其不具有感受野特性,无法做到有效的图像邻域信息交互。同时,当迁移学习数据中的空间信息与预训练数据的空间信息分布差异较大时,注意力机制中的位置编

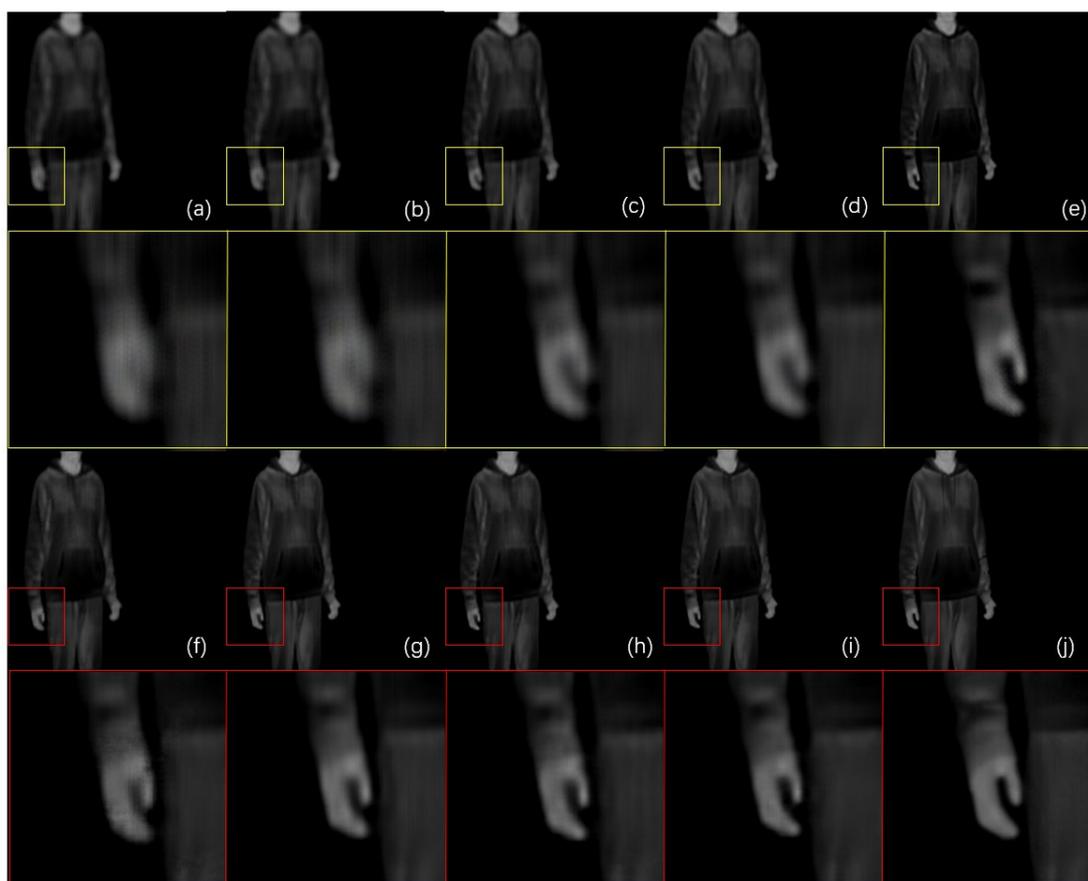


图12 不同算法超分辨图像

Fig. 12 Super resolution images based on different algorithms

表3 不同算法超分辨图像质量客观评价指标

Table 3 Objective evaluation index of super resolution image quality based on different algorithms

Different algorithm	MSE	PSNR/dB	SSIM
Nearest	25.793 6	34.015 7	0.868 3
Bicubic	27.339 4	33.762 9	0.898
SRCNN	14.322 4	36.570 6	0.901 7
SRResNet	1.489 2	46.401 3	0.924 1
SRGAN	1.592 3	46.110 6	0.956 2
ESPCNN	1.412 84	46.629 9	0.935 1
ViT	1.251 0	47.108 7	0.946 3
MDM	1.148 2	47.530 6	0.939 9
ViT with Hilbert kernel size=7	0.998 2	47.922 6	0.947 8
HR	0	$\infty$	1

码层无法有效更新,直接导致了全局建模能力的下降。针对上述问题,本文使用空间曲线填充的方法在不引入额外可学习参数的前提下,实现了利用图像特征序列对全局信息的编码,增加了额外的空间位置信息强化模型的迁移学习能力。太赫兹图像超分辨重建实验结果表明,基于本文方法编码后的图像特征序列能有效改善太赫兹超分辨图像的质量,能够获得纹理清晰、轮廓分辨能力强的结果图像。通过使用希尔伯特曲线建立的二维至一维映射,充分发掘太赫兹图像特征,有望克服空间位置编码层差异较大的问题,为不同频段图像的迁移学习与高效微调提供了有效的解决方案。在今后研究中,将对希尔伯特空间曲线填充方法的并行处理效率做进一步优化,提高其对不同波段图像性能的增强能力,改善其对不同场景、不同任务的适用性。

## References

- [1] Salih E, Selçuk S. Terahertz Technology For Military Applications [J]. *Journal of Management and Information Science*, 2015, **4**(6): 13–16.
- [2] Cheng Y Y, Qiao L B, Zhu D G, et al. Passive polarimetric imaging of millimeter and terahertz waves for personnel security screening [J]. *Optics letters*, 2021, **46**(6): 1233–1236.
- [3] Balega Y Y, Baryshev A M, Bubnov G M, et al. Superconducting Receivers for Space, Balloon, and Ground-Based Sub-Terahertz Radio Telescopes [J]. *Radiophysics and Quantum Electronics*, 2021, **2**(1): 1–22.
- [4] Xiong W H, Li L J, Ren J J, et al. Terahertz Multiple Echoes Correction and Non-Destructive Testing Based on Improved Wavelet Multi-Scale Analysis [J]. *Sensors*, 2022, **22**(9): 3477–3477.
- [5] Kiarash A. A method and system for enhancing the resolution of terahertz imaging [J]. *Measurement*, 2019, **138**: 614–619.
- [6] Wang Y, Qi F B, Wang J K. Terahertz image super-resolution based on a complex convolutional neural network [J]. *Optics letters*, 2021, **46**(13): 3123–3126.
- [7] Lu Y, Mao Q, Liu J B. Mathematical Degradation Model Learning for Terahertz Image Super-Resolution [J]. *IEEE ACCESS*, 2021, **9**: 128988–128995.
- [8] Ruan H H, Tan Z Y, Chen L T, et al. Efficient sub-pixel convolutional neural network for terahertz image super-resolution [J]. *Optics letters*, 2022, **47**(12): 3115–3118.
- [9] Yang X W, Zhang D H, Wang Z M, et al. Super-resolution reconstruction of terahertz images based on a deep-learning network with a residual channel attention mechanism [J]. *Applied optics*, 2022, **61**(12): 3363–3370.
- [10] Shi W Z, Caballero J, Huszar F, et al. Real-Time Single Image and Video Super-Resolution Using an Efficient Sub-Pixel Convolutional Neural Network [C]//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 2016: 1874–1883.
- [11] Dosovitskiy A, Lucas B, Alexander K, et al. An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale [C]// International Conference on Learning Representations. New Orleans, 2021: 1–22.
- [12] Yu Z, Hua H, Lei Z. Efficient Structure-Aware Image Smoothing by Local Extrema on Space-Filling Curve [J]. *Computer Weekly News*, 2014, **20**(9): 1253–1265.
- [13] Castro R L, Andrade D, Fraguera B B. A Winograd Minimal Filtering Algorithm Implementation in CUDA [J]. *Mathematics*, 2021, **9**(17): 2033–2033.
- [14] Jiwon K, Jung K L, Kyoung M L. Accurate Image Super-Resolution Using Very Deep Convolutional Networks [C]//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 2016: 1646–1654.
- [15] Christian L, Lucas T, Ferenc H, et al. Photo-Realistic Single Image Super-Resolution Using a Generative Adversarial Network [C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, Hawaii, USA, 2016: 105–114.
- [16] Kirillov A, Mintun E, Ravi N, et al. Segment Anything [C]//2023 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), oxford, US, 2304.0264.
- [17] Jia X, Zhu C, Li M, et al. LLVIP: A Visible-infrared Paired Dataset for Low-light Vision [C]//2021 IEEE/CVF International Conference on Computer Vision Workshops (ICCVW), Montreal, BC, Canada, 2021: 3489–3497.
- [18] Tulin O, Muhammed T, Eylul A Y, et al. Automated detection of COVID-19 cases using deep neural networks with X-ray images [J]. *Computers in Biology and Medicine*, 2020, **121**(1).
- [19] Dong C, Loy C, He K, et al. Image Super-Resolution Using Deep Convolutional Networks [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2016, **38**(2): 295–307.
- [20] He K, Zhang X, Ren S, et al. Deep Residual Learning for Image Recognition [C]//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 2016: 770–778.