

基于数据驱动的图片分辨率提升理论和技术综述

刘卫玲^{1,2}, 侯晓奎^{1,2}, 常晓明^{2*}

¹太原理工大学人工智能学院, 山西 晋中

²太原理工大学晓明研究室, 山西 太原

收稿日期: 2025年7月4日; 录用日期: 2025年10月15日; 发布日期: 2025年10月27日

摘要

在大数据与深度学习浪潮的推动下, 数据驱动模型已全面取代传统规则式算法, 成为提升图像分辨率的核心引擎。本文综述了数据驱动模型在图像分辨率提升方面的应用, 涵盖了三维重建、压缩感知、单像素成像和超分辨率技术, 并进一步探讨数据驱动模型在视觉成像、工业无损评估和医学影像处理等实际场景中的落地实践, 以及未来的发展趋势。

关键词

数据驱动模型, 图像分辨率提升, 深度学习

A Survey of Theories and Techniques for Data-Driven Image Resolution Enhancement

Weiling Liu^{1,2}, Xiaokui Hou^{1,2}, Xiaoming Chang^{2*}

¹College of Artificial Intelligence, Taiyuan University of Technology, Jinzhong Shanxi

²Xiao Ming Research Laboratory, Taiyuan University of Technology, Taiyuan Shanxi

Received: July 4, 2025; accepted: October 15, 2025; published: October 27, 2025

Abstract

Driven by the wave of big data and deep learning, data-driven models have completely replaced

*通讯作者。

文章引用: 刘卫玲, 侯晓奎, 常晓明. 基于数据驱动的图片分辨率提升理论和技术综述[J]. 嵌入式技术与智能系统, 2025, 2(3): 133-148. DOI: 10.12677/etis.2025.23011

traditional rule-based algorithms and become the core engine for improving image resolution. This paper reviews the application of data-driven models in image resolution improvement, covering 3D reconstruction, compressed sensing, single-pixel imaging, and super-resolution technology. The implementation practice of data-driven models in actual scenes such as visual imaging, industrial non-destructive evaluation, and medical image processing, and the future development trend are further discussed.

Keywords

Data-Driven Model, Image Resolution Enhancement, Deep Learning

Copyright © 2025 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

随着大数据和人工智能时代的到来,视觉数据处理技术经历了从基于规则到数据驱动的模式转变,为工业自动化、医学诊断等多个领域带来了革新。传统的分析方法依赖于手工特征和统计模型,在泛化能力、适应复杂数据分布以及真实世界成像条件下的鲁棒性方面存在局限性。而机器学习,尤其是深度学习的兴起,使得模型能够从海量数据集中自主学习层次化特征,实现了在复杂视觉分析任务中前所未有的准确性和鲁棒性。

数据驱动模型的广泛应用与计算硬件的进步、算法创新以及大规模标注数据集的可用性密切相关。早期机器学习模型如支持向量机和决策树展示了利用数据进行模式识别的潜力,但其浅层架构难以捕捉高维图像数据中的复杂空间和语义关系。卷积神经网络的兴起,通过卷积层、池化操作和非线性激活引入了层次化特征提取,开创性的工作如 AlexNet 和 ResNet 展示了 CNN 在图像分类中超越人类水平性能的能力,为其在目标检测、分割等领域的主导地位铺平了道路。

近年来,Transformer 架构的出现进一步颠覆了计算机视觉领域,其自注意力机制在建模图像中的长距离依赖和全局上下文方面表现出色,解决了 CNN 的局部性限制。视觉 Transformer (ViT)和检测 Transformer (DETR)在需要整体理解的图像描述和全景分割等任务中取得了最先进的结果。与此同时,生成模型如去噪扩散概率模型(DDPM)的出现,通过模拟扩散过程为图像重建提供了新的框架。

尽管取得了这些进展,但仍存在重大挑战。首先,退化建模不足,真实场景中的噪声、模糊类型远比合成训练数据复杂,导致模型在实际应用中表现失准;其次,计算成本高昂,高参数量的 Transformer 等架构对显存与推理时间要求极高,难以在边缘设备或实时系统中部署;第三,多模态融合薄弱,RGB、深度、红外等互补信息尚未被有效协同利用,限制了分辨率提升的上限;最后,评价指标失准,传统 PSNR/SSIM 常与主观视觉质量脱节,难以真实反映人眼对细节、纹理及整体清晰度的感知偏好。

本文聚焦图像分辨率提升的数据驱动模型,涵盖从基础技术到最新创新,系统综述三维重建、压缩感知、单像素成像及超分辨率技术。通过研究它们在视觉检测、工业无损检测(NDT)和医学成像中的应用,我们突出了这些模型的变革潜力和未解决的挑战。

本工作的贡献有三个方面:

1. 技术综合:统一传统方法(压缩感知/字典学习)与深度学习(CNN/Transformer/DDPM)的分辨率提升框架。

2. 退化 - 重建关联分析：揭示采样策略(如单像素成像)、退化模型与重建质量的耦合机制。

3. 应用场景适配：结合工业微缺陷检测、医学低剂量成像等需求提出优化路径，为研究人员和从业者提供了可操作的见解。

本文的其余部分安排如下：第 2 节专注于图像分辨率增强，包括三维重建、压缩感知和超分辨率。第 4 节讨论跨行业的应用，在第 5 节和第 6 节分别进行比较分析和未来趋势的探讨。通过这种结构，我们旨在为读者提供数据驱动视觉处理在图像分辨率提升方面的全面理解，促进理论和应用领域的创新。

2. 图像分辨率增强

2.1. 三维重建技术

三维重建技术通过结合图像采集、特征提取、匹配和模型优化等步骤，从二维图像中恢复三维场景结构，增强图像的立体感和细节。随着技术进步，该技术已发展到利用深度学习进行高效准确的重建，并在文化遗产保护、娱乐、医疗、机器人导航、工业设计和虚拟现实等多个领域得到应用。它分为基于传统多视图几何和基于深度学习的算法，正朝着深度学习、多模态融合、实时重建和交互式模型等方向进步。基于图像的三维重建技术可以分为基于传统多视图几何的三维重建算法和基于深度学习的三维重建算法。

2.1.1. 基于传统多视图几何的三维重建算法

立体视觉技术通过分析多视角图像获取深度信息以重建三维模型，分为依赖自然特征的被动方法和通过投射已知模式测量距离的主动方法。传统三维重建算法如结构光重建(SFM)和多视角立体(MVS)，通过图像配准、视差计算和特征匹配恢复三维结构，但受光照条件影响且精度有限。

从 2013 年到 2021 年，研究者们提出了多种 SFM 算法，如全局 SFM [1]、在线服务平台[2]、COLMAP 增量 SFM [3]、HSFM [4]以及基于增量 SFM 的单目三维重建方法[5]，这些技术提高了大规模三维重建的鲁棒性、精度和可扩展性。

SFM 通过特征点匹配获取相机参数，但产生稀疏点云。提供更详尽的三维信息，其中稠密点云重建 MVS 基本流程图见图 1。MVS 则通过像素级匹配生成更密集的点云，提供更详尽的三维信息。MVS 自 2006 年以来也经历了重要发展,Seitz 等人的对 MVS 算法的系统性介绍[6]、Sinha 等人的基于体素的 MVS 方法[7]，以及 Lin 等人结合双目立体视觉和特征匹配的三维重建方法[8]，推动了 MVS 领域的发展。

传统多视图几何三维重建依赖相机采集，精度不及激光点云。相机类型影响结果：彩色相机提供颜色信息但受光照影响；红外相机不受光照影响，但无法捕获颜色，实际应用时需权衡这些因素。

2.1.2. 基于深度学习的三维重建算法

传统的三维重建技术，如 SFM 和 MVS，依赖亮度一致性，在理想环境下效果良好，但在纹理弱或高反射环境下易出现不准确或空洞。深度学习方法通过编码解码过程，无需复杂校准，有效改善了这些问题。

2021 年，研究者们提出了基于深度神经网络的三维重建技术[9]-[11]，通过无监督学习、深度特征测量和增量 SFM 结构的深度融合，提高了三维重建的准确性和鲁棒性。

自 2014 年 Eigen 等人[12]首次将 CNN 应用于三维重建以来，深度学习在三维重建领域取得了显著进展，如 2015 年的多任务 CNN [13]、2017 年的 Pix2Face [14]和后续的 MVSNet 系列[15]-[20]，这些技术通过引入先进的神经网络架构，显著提升了三维重建的精度和效率。

NeRF 技术的出现推动了三维场景隐式表示的发展，从 2020 年的全连接神经网络到 2023 年的高分辨率数据训练[21]-[24]，NeRF 不断优化，提升了重建质量和细节。同时，Stucker 和 Schindler [25]、Peng 等人[26]的技术展示了深度学习在动态场景合成和多视角重建中的应用，而 Huang 等人[27]的方法则通过

二维卷积网络与三维神经网络辐射场的交互学习，为三维场景重建带来了高质量的风格化效果。这些研究证明了深度学习在解决传统三维重建挑战，如表面空洞问题方面的潜力。

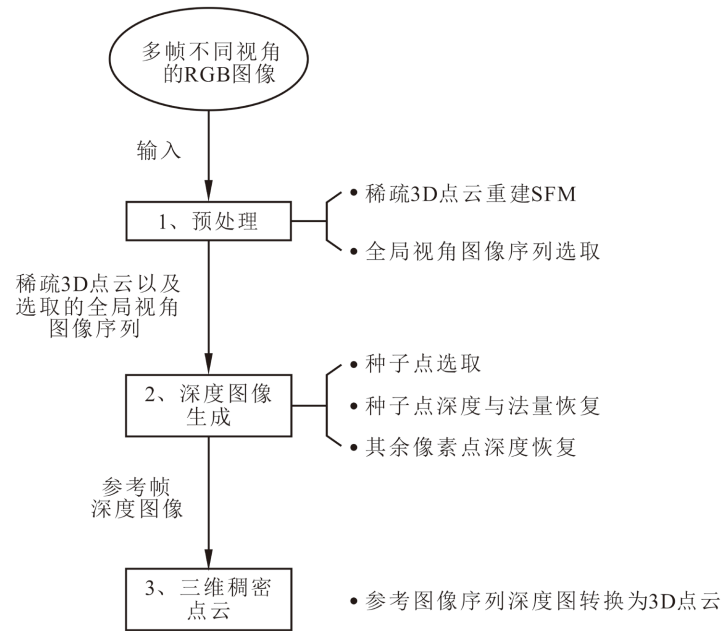


Figure 1. Basic flow chart of MVS reconstruction from dense point clouds
图 1. 稠密点云重建 MVS 基本流程图

2.1.3. 基于学习的点云稠密化方法

随着深度学习的发展，研究者提出了一系列端到端的网络，可直接由稀疏点云输出高密度表示，无需显式三维重建。2018 年的 PU-Net [28]率先用特征提取与多层感知器逐层生成新点；其后续 PU-GAN [29]、Dis-PU [30]等进一步引入生成对抗网络与几何约束，提升点的均匀性与几何一致性。另一类工作利用 NeRF 变体[31]等隐式神经表示，把点云编码为连续隐式场，让网络预测任意位置的几何概率，实现任意分辨率稠密化。为了补全缺失细节，多模态方法融合 RGB、深度或法向信息，通过跨模态注意力机制联合图像纹理与点云几何，实现高保真增强[32]。针对标注数据稀缺的现实，最新研究还探索了基于重建一致性、几何不变性的自监督或无监督生成模型，显著提高了算法在真实场景中的适应性[33]。

2.2. 压缩感知

压缩感知(Compressive Sensing, 简称 CS)是一种信号处理理论，其基本流程图见图 2，它提出了一种革命性的采样方法，允许从远低于奈奎斯特率的测量值中恢复稀疏或可压缩信号。这一理论的核心在于，如果一个信号在某种域(如时间域、空间域或频率域)中是稀疏的，那么它可以通过远少于传统采样定理要求的样本数来重建。

在图像感知与压缩领域，压缩感知的概念尤其具有吸引力。图像和视频信号通常具有内在的冗余性，这意味着它们在变换域(如小波变换或傅立叶变换)中只有少数几个系数是显著的，而其他许多系数接近于零。利用这一特性，压缩感知技术可以在保持图像质量的同时，大幅度减少需要存储或传输的数据量。例如，Yoshida 等[34]探索了将人类视觉感知引入图像压缩感知问题的可能性，通过将视觉显著性与几何特征相结合，构建视觉启发的“重要性图”，指导压缩采样过程，并结合深度图像先验(DIP)与嵌入空间流形建模(MMES)在解码端重建图像。实验结果表明，该方法在极低采样率下仍能保留关键的视觉特征，

显著优于传统的随机或均匀采样策略，为压缩感知领域带来了新的感知驱动思路。

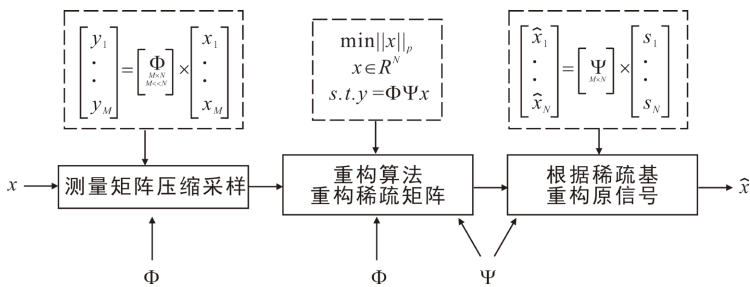


Figure 2. Basic flow chart of compressed sensing
图 2. 压缩感知基本流程图

2.2.1. 压缩感知框架

压缩感知框架是压缩感知技术的核心，它定义了如何从少量测量值中重建原始信号。压缩感知框架的核心思想是利用信号的稀疏性或可压缩性来减少数据采集量，从而实现高效的数据采集和重建。它突破了奈奎斯特采样定理的限制，允许在远低于信号带宽的速率下进行采样，并通过算法重建出高质量的信号。He 等[35]提出扩散自适应框架，通过将测量矩阵分布式存储于网络节点并引入扩散 l_0 -LMS 与 mini-batch 扩散算法，实现了稀疏信号的协同快速重建，在收敛速度与重建精度上均优于单机 l_0 -LMS。Oikonomou 等[36]提出的一种基于变分贝叶斯框架的新型压缩感知算法，实验表明该方法在多种场景下均优于现有主流算法。这些框架的成功应用，证明了压缩感知在图像重建和压缩方面的巨大潜力。

2.2.2. 字典学习

字典学习是压缩感知中一个重要的技术，它通过学习信号的稀疏表示，从而实现高效的压缩和重建。字典学习的目标是从训练数据中学习出一组基函数，使得每个信号都可以表示为这些基函数的线性组合，并且组合系数尽可能稀疏。稀疏性意味着只有少数几个系数是非零的，这有助于降低数据的存储和传输成本。S Li 等[37]提出基于稀疏编码的双字典超分辨率框架，通过 Gabor 滤波提取多尺度特征并引入残差字典补偿高频细节，显著提升了微铣刀具磨损图像的分辨率与监测精度。

2.2.3. 算法

压缩感知算法是信号重建的关键，它决定了重建的精度和效率。例如，Beck 等[38]提出了一种快速迭代收缩阈值算法(FISTA)，并将其应用于图像去模糊问题，取得了显著的性能提升，验证了 FISTA 在图像检测与压缩感知领域应用的有效性。Li 等[39]提出了一种基于总变分正则化的高效算法 TVAL3，并将其应用于单像素相机图像重建，显著提高了图像检测与压缩感知的性能，验证了 TVAL3 在图像恢复方面的有效性。表 1 为 FISTA 与 TVAL3 的相关比较。

Table 1. Relevant comparison between FISTA and TVAL3
表 1. FISTA 与 TVAL3 的相关比较

	收敛速度	内存占用	重建质量	适用数据类型
FISTA	快	低	高	稀疏信号(MRI)
TVAL3	中等	中等	极高	单像素成像

2.3. 不同图像的稀疏性

在图像处理和计算机视觉领域，稀疏性(Sparsity)是一个重要的概念，它指的是在表示图像时，图像

的某些属性或特征在某种域中只有少数几个非零元素，这些非零元素可以捕捉到图像的主要信息。图像的稀疏性是自然存在的，许多类型的图像实际上在诸如小波变换、傅立叶变换或 DCT (离散余弦变换)等正交变换下呈现高度稀疏。通过寻找最合适的稀疏表示，我们能够捕捉到图像的关键特征，同时减少冗余信息，这对于图像编码、压缩、分类和重建等方面都有积极影响，图像的稀疏表示方法分为以下几类。

2.3.1. 贪婪策略近似

贪婪策略近似方法通过迭代选择字典中与图像残差最匹配的原子，并利用最小二乘法计算原子系数，以近似求解稀疏表示问题。匹配 pursuit (MP)算法和正交匹配 pursuit (OMP)算法是最典型的贪婪策略近似方法，它们通过选择最佳匹配原子来逐步逼近图像的稀疏表示例如，Shengqin Bian 和 Lixin Zhang [40]比较了匹配追踪算法在图像重建中的应用性能，结果表明，在无噪声信号重建中，子空间追踪算法表现略优于其他算法；在二维图像重建中，当压缩比低时，子空间追踪算法也表现更佳。此外，还有许多改进的贪婪算法，如 ROMP、CoSaMP、StOMP、SP、SAMP、TMP、TBOMP、FBP 等，它们在效率、鲁棒性和收敛速度等方面进行了优化。Rajaei Khatib 等[41]提出了一种名为学习贪婪方法(LGM)的新型神经网络架构，用于稀疏编码和图像检测，该架构能够有效地学习图像的特征表示，并取得了优于传统方法的性能。

2.3.2. 约束优化策略

约束优化策略将稀疏表示问题转化为约束优化问题，并利用高效的优化方法求解。梯度投影稀疏重建(GPSR)算法将原始的无约束非光滑问题转化为一个光滑的可微约束优化问题，利用梯度下降和标准线搜索方法求解 L1 正则化问题，Li 等[42]将其与 Krylov 子空间投影相结合，应用于三维电阻层析成像(ERT)，在保持高精度图像重建的同时显著降低了计算耗时，验证了 KGPSR-BB 在实时稀疏正则化与工业过程成像中的有效性。基于内点法的稀疏表示策略，如 TNIPM 算法将 L1 范数最小化问题转化为一个无约束光滑问题，利用截断牛顿法和内点法求解 L1 正则化问题，适用于大规模稀疏表示问题。交替方向法(ADM)算法可以有效地求解 L1 正则化问题的对偶问题，进一步提高算法的效率。Yan Yang 等[43]提出了一种基于交替方向乘子法(ADMM)的深度学习架构，并将其应用于图像压缩感知，取得了优于传统方法和深度学习方法的效果，验证了深度学习在图像检测与稀疏表示方面的有效性。

2.3.3. 基于近端算法的优化策略

基于近端算法的优化策略利用近端算子迭代求解稀疏表示问题，如 ISTA、FISTA、SpaRSA、ALM 等。这些算法将 L1 正则化问题转化为 L2 最小化问题，并利用收缩算子或投影算子进行迭代求解。例如，ISTA 算法利用收缩算子迭代求解 L1 正则化问题，具有较高的收敛速度和计算效率。FISTA 算法是 ISTA 算法的改进，利用 Lipschitz 常数近似 Hessian 矩阵，并加速收敛，Li 等[44]将其与单频全息成像结合，提出 SFH-FISTA 实现 3D 稀疏毫米波成像，实验显示在 50%采样率下仍高质重建，验证了 FISTA 在雷达稀疏成像中的高效与稳健。SpaRSA 算法利用自适应连续和 BB 奇异值方法优化 L1 正则化问题，在图像压缩领域展现出高效性、灵活性、稳定性和快速收敛的优势，使其在处理大规模稀疏信号恢复问题时表现出色，尤其适用于不同类型的图像数据。ISTA、FISTA、SpaRSA 与 ALM 的相关比较见表 2。

2.3.4. 基于同伦算法的稀疏表示

基于同伦算法的稀疏表示方法利用同伦算法追踪 L1 正则化问题中参数 λ 的变化路径，逐步更新稀疏解。例如，LASSO (最小绝对收缩和选择算子)同伦算法和 BPDN (基追踪去噪)同伦算法分别追踪 LASSO 和 BPDN 问题中参数 λ 的变化路径，逐步更新稀疏解，LASSO 更多地被统计学界使用，而 BPDN 则更多地被信号处理界采用。在实际应用中，当观测数据可能包含噪声时，这两种方法都需要进行适当地调整

以处理噪声问题。此外，还有基于同伦的迭代重新加权 L1 最小化算法，它利用同伦算法更新 L1 正则化问题的权重，进一步提高算法的效率，Sining Huang 等[45]提出了一种基于扩展重加权 ℓ_1 最小化算法(ERMA)的图像恢复方法，有效提高了图像恢复的信号噪声比(SNR)、结构相似性(SSIM)和均方误差(MSE)，并通过仿真实验验证了其在图像检测和稀疏表示方面的优越性。

Table 2. Related comparison of ISTA, FISTA, SpaRSA, and ALM

表 2. ISTA、FISTA、SpaRSA、ALM 的相关比较

	收敛速度	计算复杂度	扩展性
ISTA	$O(1/k)$	低	仅适用于凸问题
FISTA	$O(1/k^2)$	低	仅适用于凸问题
SpaRSA	自适应步长加速	中低	支持非光滑问题，部分非凸扩展
ALM	超线性或线性收敛	较高	可处理非凸问题

2.4. 单像素成像

单像素成像(Single-Pixel Imaging, SPI)是一种颠覆传统的成像技术，它不依赖于成像设备上每个像素点的独立探测，而是利用单个像素探测器来捕获整个场景的图像信息。这种技术的核心在于，通过空间光调制器(Spatial Light Modulator, SLM)或者类似的设备对照明光束进行调制，产生一系列具有特定模式的照明图案，这些图案依次投射到目标物体上[46]。SPI 技术的发展，特别是在运动目标成像、盲重建、图像加密和隐藏以及照明图案优化等领域，为我们提供了新的视角和方法。

2.4.1. 运动物体成像

在运动物体成像领域，单像素成像(SPI)技术因其独特的成像机制而展现出巨大潜力，同时也面临着一系列挑战。其中最主要的挑战之一是如何在保持图像质量的同时提高成像速度。这是因为 SPI 技术在获取图像时通常需要进行多次的照明图案投射和相应的信号采集，这个过程在面对快速移动的物体时会变得尤为困难。Monin 等[47]提出一种基于循环采样矩阵与多帧运动估计的单像素成像算法，通过直接在投影域检测并补偿全局或局部运动，在目标运动过程中仍能实现高保真重建，为动态场景下的单像素成像提供了实时、稳健的解决方案。

2.4.2. 盲重建

盲重建(Blind Reconstruction)是一个在信号处理和图像分析领域常见的概念，它指的是在不知道原始信号或图像具体参数或特性的情况下，仅通过观测到的数据来恢复原始信号或图像的过程。在盲重建领域，Zhuang 等[48]提出了一种结合深度图像先验(DIP)和结构化深度神经网络的方法，用于解决盲图像去模糊(BID)问题，并在未知核大小和显著噪声的情况下表现出稳定性，验证了该方法在提高图像去模糊效果方面的有效性。Song 等[49]提出了一种基于粉噪声散斑与深度学习的计算鬼成像框架，在无需实验训练数据的情况下即可从 0.8% Nyquist 采样率中重建出高保真图像，并在未知系统响应与强噪声条件下实现鲁棒盲重建，验证了该方法在极低采样与复杂环境下的有效性。

2.4.3. 图像加密和隐藏

SPI 技术在图像加密和隐藏方面利用其独特的成像机制提供了新的安全策略，它通过随机相位掩模对图像进行编码，生成难以识别的噪声状图案，从而保护图像内容不被未经授权访问。此外，SPI 技术结合混沌理论可以生成高度随机的编码模式，进一步增强图像的安全性。Zhang 等[50]提出了一种基于混沌棕

桐相位掩模(CPPM)和菲涅耳变换(FrT)的光学单通道彩色图像加密方案,并将其应用于光学信息安全领域,取得了显著的安全性提升,验证了该方案在提取光学图像隐藏信息方面的有效性,这为 SPI 在图像加密领域提供了新的视角。

2.5. 人工智能超分辨率及其局限性

超分辨率技术(Super-Resolution, 简称 SR)是一种旨在提高图像或视频分辨率的技术。传统的图像放大方法往往会导致图像模糊和失真,而超分辨率技术则通过算法重建图像的细节,生成比原始图像更清晰、分辨率更高的图像,同时保留原始内容和结构,Liu 等[51]提出了一种基于 Cycle-GAN 的超编码分辨率重建方法,通过无配对训练策略在远低于奈奎斯特采样条件下实现 $2\times$ 超分辨成像,实验验证其在 $3.125\%\sim 25\%$ 采样率下显著提升图像细节与边缘锐度,为低数据量、高保真超分辨率成像提供了新途径。超分辨率技术与人工智能结合的关键优势在于其能够从大量的数据中学习到复杂的特征,并利用这些特征来增强图像的细节和质量。这种结合不仅提高了图像的质量,还扩展了超分辨率技术在多个领域的应用范围,为图像处理领域带来了新的可能性。

2.5.1. 单一图像超分辨率

单一图像超分辨率(SISR)技术是一种图像增强方法,它旨在从单个低分辨率图像中恢复出高分辨率图像。SISR 的挑战在于,由于成像系统的局限性或数据获取过程中的约束,低分辨率图像丢失了高频细节信息。为了克服这些限制,SISR 通常采用先进的算法,如基于学习的方法,特别是深度学习技术。Liang 等[52]提出了一种基于 Swin Transformer 的端到端图像复原框架 SwinIR,通过局部-全局混合注意力机制直接学习从低分辨率到高分辨率的映射,在多个超分辨率任务及退化场景下均显著优于现有 CNN 方法,验证了 Transformer(结构示意图见 图 3)在单图像超分辨率中的有效性与高效性。Zhang 等[53]提出了一种由随机模糊-降采样-噪声级联并可随机洗牌的实用退化模型,联合 ESRGAN 架构端到端训练出 BSRGAN,在未知复杂退化的真实图像上实现盲超分辨率,显著提升了视觉质量与鲁棒性,验证了该方法在实际场景中的有效性。

2.5.2. 多图像超分辨率

多图像超分辨率(MISR)技术利用多个低分辨率图像重建一个高分辨率图像。这种方法通常涉及到图像配准,以确保多个图像中的场景对齐,然后通过融合技术合并图像信息以提高分辨率。与单图像超分辨率(SISR)技术不同,MISR 通过整合多幅图像中的互补信息,能够更有效地恢复高分辨率图像的细节和结构。SISR 与 MISR 的相关比较见表 3。

Salvetti 等[54]提出了一种基于残差特征注意力的深度神经网络 RAMS,利用 3D 卷积同时融合多幅低分辨率遥感影像的时空信息,实现 $3\times$ 超分辨率重建,在公开 Proba-V 数据集上显著优于单图与现有多图方法,验证了其在大规模遥感场景中的有效性与可迁移性。MISR 的优势在于它可以利用多个视角中的冗余信息来增强细节并减少噪声,这种方法特别适合于场景相对静态的情况,Xiu 等[55]提出的一种新的端到端网络结构 CoT-MISR,结合了卷积和 Transformer 的优势,有效利用低分辨率图像的局部和全局信息,在 PROBA-V 数据集上取得了目前多图像超分辨率任务的最佳性能,为遥感图像融合提供了新的思路。随着计算能力的提升和算法的改进,MISR 技术在处理大型图像数据集和提供更高分辨率图像方面变得更加有效。

2.5.3. 局限性

当前人工智能与图像超分辨率的结合取得了一定进展,但仍面临挑战。首先,虽然 CNN 模型被广泛使用,但 GAN 模型在处理复杂场景和不同缩放任务时更具优势,未来的研究需要探索如何结合两者的优

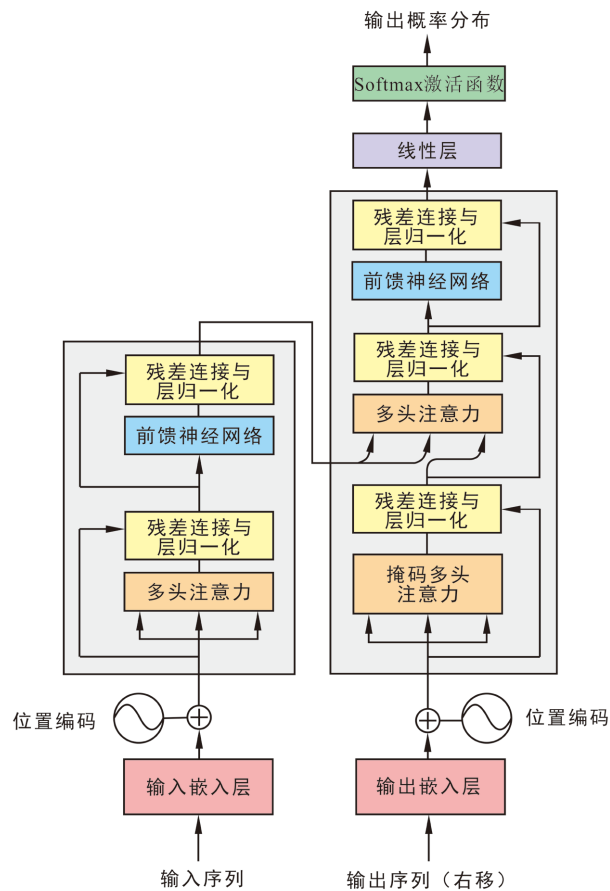


Figure 3. Schematic diagram of the structure of Transformers
图 3. Transformers 结构示意图

Table 3. Correlation comparison between SISR and MISR

表 3. SISR 与 MISR 的相关比较

	输入类型	重建质量	计算成本	典型算法
SISR	单张低分辨率图	中等	低	SRCNN、ESRGAN
MISR	多张低分辨率图	高	高	CoT-MISR、MVSR

势并简化网络结构[56]。其次，现有模型在适应多变的实际场景方面存在困难，需要设计更灵活的模型来利用不同场景的先验知识。此外，可以引入深度学习中的多种学习方法，如注意力机制和多因素学习，以提升模型性能。在图像退化和采样方法方面，需要探索更接近真实情况的模型和方法，以弥补现有方法与实际应用之间的差距。最后，现有的目标函数和评估标准存在一定的局限性，需要开发更合理的损失函数和质量评价方法，以平衡模型的精度和感知质量。

3. 应用案例

3.1. 视觉检测

视觉检测是利用计算机视觉技术来识别、定位和分析图像或视频中的物体和事件。它不仅仅是一个技术过程，更是一种使机器能够理解和解释视觉信息的革命性能力，这种能力使得机器能够在各种复杂

的环境中执行任务。数据驱动模型，尤其是深度学习模型，已经成为实现这一能力的核心工具。这些模型通过从大量图像数据中学习，能够自动提取特征并建立复杂的模式识别系统。深度学习的发展，特别是卷积神经网络(CNN)的广泛应用，已经极大地推动了视觉检测技术的进步。CNN 在图像和视频分析任务中展现出了前所未有的性能，使得机器能够在没有人类直接干预的情况下，准确地识别和定位图像中的物体。

随着技术的不断进步，视觉检测的应用已经扩展到了许多新的领域。在馆藏数字化领域，倪劼等[57]提出一种改进的 Real-ESRGAN 模型，针对馆藏近代低分辨率图像引入多尺度特征融合与通道注意力机制，在 4×超分辨率重建后将图像送入后续视觉检测流程，显著提升了历史文献中文字、图案等关键细节的检出率(PSNR ↑ 3 dB, SSIM ↑ 0.0672)，为图书馆数字化展示与智能检索提供了高保真视觉基础。在体育分析中，视觉检测技术被用来追踪运动员表现和比赛动态，谢竞光和程新年[58]提出的一个结合生成对抗网络(GANs)和循环神经网络(RNNs)的系统，用于预测排球比赛中的战术数据。该系统利用计算机视觉技术自动从比赛视频中提取关键信息，并运用深度学习进行智能分析。这一方法能有效识别和理解排球比赛中的复杂战术行为，并提供准确预测，为战术分析、训练和比赛决策提供技术支持。

尽管视觉检测技术已经取得了显著的成就，但仍有许多挑战需要克服。例如，如何在不同的光照条件和复杂背景中保持高准确性，如何提高模型的实时处理能力，以及如何处理大规模图像数据集。未来的研究将继续探索更高效的算法，提高模型的泛化能力和解释性，以及开发新的硬件和软件解决方案，以满足不断增长的应用需求。

3.2. 工业无损检测

在工业应用中，无损检测(NDT)扮演着至关重要的角色，它确保了产品在不遭受物理损伤的情况下，其质量和安全性得到有效保障。随着与图像检测技术的融合，NDT 的能力和效率得到了显著提升，成为了维护工业产品质量的强有力工具。利用高分辨率成像技术、尖端图像处理算法以及机器学习模型，这些综合技术能够自动化地识别材料表面的微观缺陷，精确评估焊接与连接的质量，持续监测复合材料的结构完整性，并准确验证电子组件的制造精度。

在这一领域，高分辨率成像技术的应用尤为关键，其提供的超高像素密度和细节还原能力，使得检测过程能够捕获材料表面的微观结构。通过提升图像的空间分辨率与对比度，这些技术为后续分析提供了更清晰、更丰富的视觉数据基础，使微米级甚至纳米级的特征可视化成为可能，从而为无损检测的精确性奠定技术前提。

随着图像检测技术的不断发展，其在工业无损检测领域的应用越来越广泛。在窄间隙焊缝的缺陷检测中，Nicolson 等[59]以双串联相控阵超声与 FMC 成像，在窄间隙焊缝中实时分辨亚毫米级未熔合缺陷，为核电厚壁焊接提供高分辨率在线检测，使核电与可再生能源等领域厚壁结构的高质量、低成本制造取得可靠保障。在涡轮叶片 X 射线成像领域，马钟、赵歆波等人[60]提出了一种基于频域亚像素配准与非均匀插值的超分辨率重建技术，该技术利用多幅含相对位移的低分辨率 DR 图像，通过傅里叶域精确配准和双调和样条插值，实现了分辨率提升至原图 4 倍的高保真成像，有效揭示了叶片表面及内部的细微缺陷，实验验证其在航空无损检测中具有显著的工程应用价值。

这些研究表明，高分辨率图像技术的进步在工业无损检测领域正发挥越来越关键的作用。通过持续提升成像设备的解析能力与信噪比，该技术为材料微观结构的可视化提供了更强大的支撑，从而推动无损检测向更高精度和可靠性发展。

3.3. 医疗成像

医疗成像技术是一种用于获取人体内部结构图像的非侵入性或微创性技术，它在临床诊断和治疗中

发挥着至关重要的作用。通过不同的成像原理,如 X 射线、超声波、磁共振、放射性核素等,医疗成像技术通过不断提升空间分辨率和对比度,能够以微米级精度呈现人体内部器官、组织的微观结构,为临床研究提供高保真图像基础。数据驱动模型显著优化了成像设备的解析能力与信噪比,使纳米级生物特征的可视化成为可能,推动医疗成像向更高清晰度和效率发展。

医疗成像技术结合图像检测旨在提高疾病诊断的准确性和效率,高分辨率成像技术(如 CT、MRI)通过提升像素密度和层析精度,为图像处理算法提供了更丰富的结构细节。基于深度学习的超分辨率重建和噪声抑制模型,能进一步优化原始图像的纹理清晰度与边界锐度,使细胞级结构或微血管形态等亚视觉特征得以清晰呈现,为医学研究奠定高质量数据基础。

Sobek 等[61]开发了一个名为 Med-YOLO 的三维医学图像目标检测框架,该框架基于 YOLO 模型。Med-YOLO 通过 3D 版本替换了 2D 神经网络层,使其能够理解和分析医学图像中的三维结构。Zhe Guo 等[62]的多模态分割算法,利用超高分辨率图像叠加,实现了亚毫米级软组织纹理的精准分层映射。此外,AI 医学影像模型如 SLiViT [63],它作为一种深度学习模型,能够快速高效地进行专家级图像分析,该模型通过融合多尺度分辨率数据,在低信噪比条件下仍能保持生物标记物成像的完整性,验证了高分辨率技术对复杂医学图像分析的普适价值。这些技术的应用,使得医学影像检测更加精确,有助于提高疾病诊断的准确性和治疗的成功率。

医疗成像中的图像检测技术面临的挑战在于平衡辐射剂量与分辨率需求(如低剂量 CT)、突破衍射极限的微观成像,以及海量高分辨率数据的实时处理。未来研究将聚焦开发轻量化超分辨率算法、量子成像传感器等硬件革新,以突破现有分辨率极限,实现无损活体纳米级成像。

4. 比较与讨论

在图像检测领域,传统图像处理方法(如 SIFT、HOG 特征提取)依赖人工设计特征与统计模型(PCA、LDA),其优势在于算法透明、计算效率高,但在复杂场景(如动态模糊、低纹理区域)中泛化能力有限。而深度学习模型(CNN、Transformer)通过数据驱动的端到端学习,自动挖掘多层次特征表达,显著提升了对噪声、形变等干扰的鲁棒性。

不同领域对图像质量的差异化需求,正推动分辨率增强技术沿着“场景定制”路径快速演进:在工业检测中,时序生成模型通过捕捉视频帧间动态信息,显著改善了运动模糊图像的恢复效果;在医学成像领域,三维分割算法借助超高分辨率数据,突破了压缩感知稀疏重建的精度瓶颈;而在文化遗产保护场景,改进的超分辨率模型针对古籍褪色文字进行纹理保真优化,有效避免了传统方法对模糊字形的误判。

针对视觉质量优化问题,当前研究方法呈现出明显的技术分界:传统算法凭借人工设计特征的高可解释性及低计算复杂度占据基础优势,但其表征能力受限于先验模型构建范式,在动态模糊、弱纹理等复杂成像条件下的泛化性能呈现断崖式衰减;而深度学习方法通过端到端的层次化特征学习机制显著提升了系统的鲁棒性,却伴随着模型参数量激增引发的计算资源消耗与训练数据需求的同步激增。为此,现代分辨率增强技术已突破传统“同质化”超分范式,创新性地构建面向工业检测、医学影像、文化遗产修复等垂直领域的自适应增强框架。通过引入时序生成对抗网络建模动态退化过程、构建三维点云高密度重建的物理约束模型以及融合多尺度纹理先验的损失函数,该技术范式成功实现了从低层次像素级复现到高层次语义信息增强的技术跃迁,为不同应用场景下的图像质量提升提供了理论依据与方法支撑。

5. 发展趋势

基于数据驱动模型的现有突破与共性挑战,未来研究将聚焦以下方向:

(1) 提升数据预处理和特征提取的精度

在视觉数据的预处理阶段，对图像进行降噪、增强和归一化等操作至关重要。这些操作能够提高后续模型训练的效率和检测的准确性。特征提取作为视觉数据处理的关键步骤，直接影响到模型的性能。未来的研究需要开发更先进的算法，以提高特征提取的精度和鲁棒性，尤其是在复杂环境下对目标的识别和分类。例如，深度学习模型如卷积神经网络(CNN)和 Transformer 在特征提取方面展现出了强大的能力，未来的研究可以进一步探索这些模型在视觉数据处理中的应用。

(2) 探索基于深度学习的多模态融合技术。

多模态融合技术通过整合来自不同传感器或不同来源的数据，能够提供更全面的图像信息，提高检测的准确性和鲁棒性。深度学习模型，尤其是多模态深度学习模型，为处理和融合多源数据提供了强大的工具。未来的研究可以探索如何利用深度学习模型来处理和融合来自不同模态的数据，以提高视觉数据处理的性能。

(3) 发展小样本学习和模型可解释性

在实际应用中，某些场景下标注数据的获取可能非常昂贵或不可行，这就需要模型能够在少量标注数据上进行有效的学习。小样本学习技术可以帮助模型在数据稀缺的情况下进行学习。同时，模型的可解释性也是视觉数据处理领域的一个重要研究方向，它可以帮助用户理解模型的决策过程，增强对模型的信任。

(4) 推动无监督和自监督学习技术的发展

无监督学习技术可以在没有标注数据的情况下发现数据中的模式和结构，这对于大规模图像数据的处理尤为重要。自监督学习技术通过设计预测任务，使模型能够从未标注的数据中学习有用的特征表示。这些技术的发展将进一步拓宽视觉数据处理技术的应用范围。

(5) 应对实际成像环境的复杂性

实际成像环境的复杂性对视觉数据处理技术提出了挑战，包括光照变化、遮挡、动态场景等问题。未来的研究需要开发更加鲁棒的模型，以应对这些复杂环境的影响。

综上所述，视觉数据处理领域的驱动模型正朝着提高预处理和特征提取精度、探索多模态融合技术、发展小样本学习和模型可解释性、推动无监督和自监督学习技术发展以及应对实际成像环境复杂性等方向发展。这些趋势将共同推动视觉数据处理技术的进步，以适应不断增长的实际应用需求。

6. 总结与展望

本文系统综述了视觉数据处理数据驱动模型的发展脉络与技术体系，重点聚焦图像分辨率提升的关键技术路径与应用前景。通过对三维重建、压缩感知、单像素成像及超分辨率等核心技术的剖析，揭示了数据驱动模型从传统机器学习到深度学习(CNN、Transformer、DDPM)的范式演进。研究表明，物理机制与数据驱动的深度融合、多模态协同重建、边缘轻量化部署已成为突破现有技术瓶颈的关键方向。综上所述，笔者认为视觉数据处理领域下一步的研究重点方向如下。

(1) 多模态融合技术的发展：目前的融合框架多数是基于两种不同来源的图像，未来研究应探索将更多种类的图像数据融合在一起，以期获得更丰富全面的特征信息。这将有效改善当前大部分融合算法提取的特征为单一特征的局限性，提高视觉数据处理的准确性和鲁棒性。

(2) 预处理技术的整合与智能化：随着图像配准、特征提取等预处理技术的快速发展，未来的研究应着力于将这些处理模块与融合模块整合为一套智能的处理系统，实现一站式融合。这将降低对输入图像的要求，扩大多源图像融合技术的应用范围，并提升用户体验。

(3) 评价指标的标准化：融合算法的性能并不总是与融合图像性能完全匹配，现有的评价指标也不完

全等同于图像的主观评价。因此,为了更客观地评价融合效果,未来研究需要制定和规范评价标准,提升评价质量,以更好地指导图像融合技术的发展和应。

综上所述,视觉数据处理领域的驱动模型正面临着新的挑战和机遇,未来的研究将在多模态融合技术、预处理技术的整合与智能化、评价指标的标准化等方面取得新的进展,推动视觉数据处理技术向更高精度、更广应用和更深理解的方向发展。

致 谢

衷心感谢山西省青年科学研究项目“高温环境下炉外磁场差分实现大型矿热炉电极端部位置的在线检测”(202103021223067)所提供的支持。

参考文献

- [1] Moulon, P., Monasse, P. and Marlet, R. (2013) Global Fusion of Relative Motions for Robust, Accurate and Scalable Structure from Motion. 2013 *IEEE International Conference on Computer Vision*, Sydney, 1-8 December 2013, 3248-3255. <https://doi.org/10.1109/iccv.2013.403>
- [2] Heller, J., Havlena, M., Jancosek, M., Torii, A. and Pajdla, T. (2015) 3D Reconstruction from Photographs by CMP SfM Web Service. 2015 *14th IAPR International Conference on Machine Vision Applications (MVA)*, Tokyo, 18-22 May 2015, 30-34. <https://doi.org/10.1109/mva.2015.7153126>
- [3] Schonberger, J.L. and Frahm, J. (2016) Structure-from-Motion Revisited. 2016 *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, 27-30 June 2016, 4104-4113. <https://doi.org/10.1109/cvpr.2016.445>
- [4] Cui, H., Gao, X., Shen, S. and Hu, Z. (2017) HSFM: Hybrid Structure-from-Motion. 2017 *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Honolulu, 21-26 July 2017, 2393-2402. <https://doi.org/10.1109/cvpr.2017.257>
- [5] Yin, H.Y. and Yu, H.Y. (2020) Incremental SFM 3D Reconstruction Based on Monocular. 2020 *13th International Symposium on Computational Intelligence and Design (ISCID)*, Hangzhou, 12-13 December 2020, 17-21. <https://doi.org/10.1109/iscid51228.2020.00011>
- [6] Seitz, S.M., Curless, B., Diebel, J., Scharstein, D. and Szeliski, R. (2006) A Comparison and Evaluation of Multi-View Stereo Reconstruction Algorithms. 2006 *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, New York, 17-22 June 2006, 519-528.
- [7] Sinha, S.N., Mordohai, P. and Pollefeys, M. (2007) Multi-View Stereo via Graph Cuts on the Dual of an Adaptive Tetrahedral Mesh. 2007 *IEEE 11th International Conference on Computer Vision*, Rio de Janeiro, 14-21 October 2007, 1-8. <https://doi.org/10.1109/iccv.2007.4408997>
- [8] Lin, X.B., Wang, J.X. and Lin, C. (2020) Research on 3D Reconstruction in Binocular Stereo Vision Based on Feature Point Matching Method. 2020 *IEEE 3rd International Conference on Information Systems and Computer Aided Education (ICISCAE)*, Dalian, 27-29 September 2020, 551-556. <https://doi.org/10.1109/iciscae51034.2020.9236889>
- [9] Wang, Y.X., Lu, Y.W., Xie, Z.H. and Lu, G.Y. (2021) Deep Unsupervised 3D SfM Face Reconstruction Based on Massive Landmark Bundle Adjustment. *Proceedings of the 29th ACM International Conference on Multimedia*, Chengdu, 20-24 October 2021, 1350-1358. <https://doi.org/10.1145/3474085.3475689>
- [10] Lindenberger, P., Sarlin, P., Larsson, V. and Pollefeys, M. (2021) Pixel-Perfect Structure-From-Motion with Feature-metric Refinement. 2021 *IEEE/CVF International Conference on Computer Vision (ICCV)*, Montreal, 10-17 October 2021, 5967-5977. <https://doi.org/10.1109/iccv48922.2021.00593>
- [11] Zhou, L., Zhang, Z., Jiang, H., Sun, H., Bao, H. and Zhang, G. (2021) DP-MVS: Detail Preserving Multi-View Surface Reconstruction of Large-Scale Scenes. *Remote Sensing*, **13**, Article 4569. <https://doi.org/10.3390/rs13224569>
- [12] Eigen D., Puhersch, C. and Fergus, R. (2014) Depth Map Prediction from a Single Image Using a Multi-Scale Deep Network. *International Conference on Neural Information Processing Systems*, Cambridge, 8-13 December 2014, 2366-2374.
- [13] Eigen, D. and Fergus, R. (2015) Predicting Depth, Surface Normals and Semantic Labels with a Common Multi-Scale Convolutional Architecture. 2015 *IEEE International Conference on Computer Vision (ICCV)*, Santiago, 7-13 December 2015, 2650-2658. <https://doi.org/10.1109/iccv.2015.304>
- [14] Crispell, D. and Bazik, M. (2017) Pix2Face: Direct 3D Face Model Estimation. 2017 *IEEE International Conference on Computer Vision Workshops (ICCVW)*, Venice, 22-29 October 2017, 2512-2518. <https://doi.org/10.1109/iccvw.2017.295>
- [15] Yao, Y., Luo, Z., Li, S., Fang, T. and Quan, L. (2018) MVSNet: Depth Inference for Unstructured Multi-View Stereo.

- In: *Lecture Notes in Computer Science*, Springer, 785-801. https://doi.org/10.1007/978-3-030-01237-3_47
- [16] Yao, Y., Luo, Z., Li, S., Shen, T., Fang, T. and Quan, L. (2019) Recurrent MVSNet for High-Resolution Multi-View Stereo Depth Inference. 2019 *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Long Beach, 15-20 June 2019, 5520-5529. <https://doi.org/10.1109/cvpr.2019.00567>
 - [17] Chen, R., Han, S., Xu, J. and Su, H. (2019) Point-Based Multi-View Stereo Network. 2019 *IEEE/CVF International Conference on Computer Vision (ICCV)*, Seoul, 27 October-2 November 2019, 1538-1547. <https://doi.org/10.1109/iccv.2019.00162>
 - [18] Zhang, J., Yao, Y., Li, S., Luo, Z. and Fang, T. (2020) Visibility-Aware Multi-View Stereo Network. *Proceedings of the British Machine Vision Conference 2020*, Manchester, 7-10 September 2020, 184-200. <https://doi.org/10.5244/c.34.109>
 - [19] Wei, Z., Zhu, Q., Min, C., Chen, Y. and Wang, G. (2021) AA-RMVSNet: Adaptive Aggregation Recurrent Multi-View Stereo Network. 2021 *IEEE/CVF International Conference on Computer Vision (ICCV)*, Montreal, 10-17 October 2021, 6167-6176. <https://doi.org/10.1109/iccv48922.2021.00613>
 - [20] Peng, R., Wang, R., Wang, Z., Lai, Y. and Wang, R. (2022) Rethinking Depth Estimation for Multi-View Stereo: A Unified Representation. 2022 *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, New Orleans, 18-24 June 2022, 18-24. <https://doi.org/10.1109/cvpr52688.2022.00845>
 - [21] Mildenhall, B., Srinivasan, P.P., Tancik, M., Barron, J.T., Ramamoorthi, R. and Ng, R. (2020) NeRF: Representing Scenes as Neural Radiance Fields for View Synthesis. In: *Lecture Notes in Computer Science*, Springer, 405-421. https://doi.org/10.1007/978-3-030-58452-8_24
 - [22] Yen-Chen, L., Florence, P., Barron, J.T., Rodriguez, A., Isola, P. and Lin, T. (2021). iNeRF: Inverting Neural Radiance Fields for Pose Estimation. 2021 *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, Prague, 27 September-1 October 2021, 1323-1330. <https://doi.org/10.1109/iros51168.2021.9636708>
 - [23] Xu, Q.G., Xu, Z., Philip, J., Bi, S., Shu, Z., Sunkavalli, K., et al. (2022) Point-NeRF: Point-Based Neural Radiance Fields. 2022 *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, New Orleans, 18-24 June 2022, 5428-5438. <https://doi.org/10.1109/cvpr52688.2022.00536>
 - [24] Xu, L., Xiangli, Y., Peng, S., Pan, X., Zhao, N., Theobalt, C., et al. (2023) Grid-Guided Neural Radiance Fields for Large Urban Scenes. 2023 *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Vancouver, 17-24 June 2023, 8296-8306. <https://doi.org/10.1109/cvpr52729.2023.00802>
 - [25] Stucker, C. and Schindler, K. (2020) ResDepth: Learned Residual Stereo Reconstruction. 2020 *IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, Seattle, 14-19 June 2020, 707-716. <https://doi.org/10.1109/cvprw50498.2020.00100>
 - [26] Peng, S.D., Zhang, Y.Q., Xu, Y.H., Wang, Q.Q., Shuai, Q., Bao, H.J. and Zhou, X.W. (2021) Neural Body: Implicit Neural Representations with Structured Latent Codes for Novel View Synthesis of Dynamic Humans. 2021 *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Nashville, 20-25 June 2021, 9050-9059. <https://doi.org/10.1109/cvpr46437.2021.00894>
 - [27] Huang, Y.H., He, Y., Yuan, Y.J., Lai, Y.K. and Gao, L. (2022) StylizedNeRF: Consistent 3D Scene Stylization as Stylized Nerf via 2D-3D Mutual Learning. 2022 *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, New Orleans, 18-24 June 2022, 18321-18331. <https://doi.org/10.1109/cvpr52688.2022.01780>
 - [28] Yu, L., Li, X., Fu, C., Cohen-Or, D. and Heng, P. (2018) Pu-Net: Point Cloud Upsampling Network. 2018 *IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Salt Lake City, 18-23 June 2018, 2790-2799. <https://doi.org/10.1109/cvpr.2018.00295>
 - [29] Li, R., Li, X., Fu, C., Cohen-Or, D. and Heng, P. (2019) PU-GAN: A Point Cloud Upsampling Adversarial Network. 2019 *IEEE/CVF International Conference on Computer Vision (ICCV)*, Seoul, 27 October-2 November 2019, 7202-7211. <https://doi.org/10.1109/iccv.2019.00730>
 - [30] He, Y., Tang, D., Zhang, Y., Xue, X. and Fu, Y. (2023) Grad-Pu: Arbitrary-Scale Point Cloud Upsampling via Gradient Descent with Learned Distance Functions. 2023 *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Vancouver, 17-24 June 2023, 5354-5363. <https://doi.org/10.1109/cvpr52729.2023.00518>
 - [31] Mildenhall, B., Srinivasan, P.P., Tancik, M., Barron, J.T., Ramamoorthi, R. and Ng, R. (2020) NeRF: Representing Scenes as Neural Radiance Fields for View Synthesis. In: *Lecture Notes in Computer Science*, Springer, 405-421. https://doi.org/10.1007/978-3-030-58452-8_24
 - [32] Li, S.Y., Yang, W. and Liao, Q. (2024) PMAFusion: Projection-Based Multi-Modal Alignment for 3D Semantic Occupancy Prediction. 2024 *IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, Seattle, 17-18 June 2024, 3627-3634. <https://doi.org/10.1109/cvprw63382.2024.00366>
 - [33] Jiang, H.W., Huang, Q. and Pavlakos, G. (2024) Real3D: Scaling up Large Reconstruction Models with Real-World Images. <https://doi.org/10.48550/arXiv.2406.08479>
 - [34] Yoshida, S., Sun, Z., Yoshizawa, S., Michikawa, T., Noda, S., Micheletto, R., et al. (2024). Image Compressed Sensing

- Based on Vision-Inspired Importance Maps. 2024 *IEEE International Conference on Imaging Systems and Techniques (IST)*, Tokyo, 14-16 October 2024, 1-6. <https://doi.org/10.1109/ist63414.2024.10759151>
- [35] He, Y.C., Wang, F., Wang, S.Y. and Chen, B.D. (2017) Diffusion Adaptation Framework for Compressive Sensing Reconstruction. *Signal Processing*, **176**, Article 107660.
- [36] Oikonomou, V.P., Nikolopoulos, S. and Kompatsiaris, I. (2019) A Novel Compressive Sensing Scheme under the Variational Bayesian Framework. 2019 *27th European Signal Processing Conference (EUSIPCO)*, A Coruna, 2-6 September 2019, 1-5. <https://doi.org/10.23919/eusipco.2019.8902704>
- [37] Li, S., Ling, Z. and Zhu, K. (2024) Image Super Resolution by Double Dictionary Learning and Its Application to Tool Wear Monitoring in Micro Milling. *Mechanical Systems and Signal Processing*, **206**, Article 110917. <https://doi.org/10.1016/j.ymssp.2023.110917>
- [38] Beck, A. and Teboulle, M. (2009) A Fast Iterative Shrinkage-Thresholding Algorithm with Application to Wavelet-Based Image Deblurring. 2009 *IEEE International Conference on Acoustics, Speech and Signal Processing*, Taipei, 19-24 April 2009, 693-696. <https://doi.org/10.1109/icassp.2009.4959678>
- [39] Li, C.B. (2010) An Efficient Algorithm for Total Variation Regularization with Applications to the Single Pixel Camera and Compressive Sensing. Master's Thesis, Rice University.
- [40] Bian, S. and Zhang, L. (2021) Overview of Match Pursuit Algorithms and Application Comparison in Image Reconstruction. 2021 *IEEE Asia-Pacific Conference on Image Processing, Electronics and Computers (IPEC)*, Dalian, 14-16 April 2021, 216-221. <https://doi.org/10.1109/ipec51340.2021.9421295>
- [41] Khatib, R., Simon, D. and Elad, M. (2020) Learned Greedy Method (LGM): A Novel Neural Architecture for Sparse Coding and beyond. *Journal of Visual Communication and Image Representation*, **77**, Article 103095. <https://doi.org/10.1016/j.jvcir.2021.103095>
- [42] Li, S., Wang, H., Liu, T., Cui, Z., Chen, J.N. and Xia, Z. (2021) A Fast Barzilai-Borwein Gradient Projection for Sparse Reconstruction Algorithm Based on 3D Modeling: Application to ERT Imaging. *IEEE Access*, **9**, 152913-152922. <https://doi.org/10.1109/access.2021.3127695>
- [43] Yang, Y., Sun, J., Li, H. and Xu, Z. (2020) ADMM-CSNet: A Deep Learning Approach for Image Compressive Sensing. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **42**, 521-538. <https://doi.org/10.1109/tpami.2018.2883941>
- [44] Li, X., Ran, J. and Zhou, Z. (2022) An Efficient 3D Radar Imaging Algorithm Based on FISTA. 2022 *IEEE 9th International Symposium on Microwave, Antenna, Propagation and EMC Technologies for Wireless Communications (MAPE)*, Chengdu, 26-29 August 2022, 419-423. <https://doi.org/10.1109/mape53743.2022.9935219>
- [45] Huang, S., Chen, Y. and Qiao, T. (2021) An Extended Reweighted ℓ_1 Minimization Algorithm for Image Restoration. *Mathematics*, **9**, Article 3224. <https://doi.org/10.3390/math9243224>
- [46] Edgar, M.P., Gibson, G.M. and Padgett, M.J. (2018) Principles and Prospects for Single-Pixel Imaging. *Nature Photonics*, **13**, 13-20. <https://doi.org/10.1038/s41566-018-0300-7>
- [47] Monin, S., Hahamovich, E. and Rosenthal, A. (2021) Single-Pixel Imaging of Dynamic Objects Using Multi-Frame Motion Estimation. *Scientific Reports*, **11**, Article No. 7712. <https://doi.org/10.1038/s41598-021-83810-z>
- [48] Zhuang, Z., Li, T.H., Wang, H.K., et al. (2022) Blind Image Deblurring with Unknown Kernel Size and Substantial Noise. <https://doi.org/10.48550/arXiv.2208.09483>
- [49] Song, H., Nie, X., Su, H., Chen, H., Zhou, Y., Zhao, X., et al. (2021) 0.8% Nyquist Computational Ghost Imaging via Non-Experimental Deep Learning. *Optics Communications*, **520**, Article 128450. <https://doi.org/10.1016/j.optcom.2022.128450>
- [50] Zhang, H., Zhao, Q., Xu, W., Wang, Y., Li, F., Liu, S., et al. (2024) Optical Single-Channel Color Image Encryption Based on Chaotic Palmprint Phase Masks. *Journal of Optics*, **53**, 3342-3350. <https://doi.org/10.1007/s12596-023-01510-5>
- [51] Liu, S.P., Wu, H., Li, Q., Meng, X. and Yin, Y. (2023) Super-Coding Resolution Single-Pixel Imaging Based on Unpaired Data-Driven Deep Learning. *Optics and Lasers in Engineering*, **170**, Article 107786. <https://doi.org/10.1016/j.optlaseng.2023.107786>
- [52] Liang, J.Y., Cao, J.Z., Sun, G.L., Zhang, K., et al. (2021) SwinIR: Image Restoration Using Swin Transformer. 2021 *IEEE/CVF International Conference on Computer Vision Workshops (ICCVW)*, Montreal, 11-17 October 2021, 1833-1844. <https://doi.org/10.1109/iccvw54120.2021.00210>
- [53] Zhang, K., Liang, J.Y., Van Gool, L. and Timofte, R. (2021) Designing a Practical Degradation Model for Deep Blind Image Super-Resolution. 2021 *IEEE/CVF International Conference on Computer Vision (ICCV)*, Montreal, 10-17 October 2021, 4771-4780. <https://doi.org/10.1109/iccv48922.2021.00475>
- [54] Salvetti, F., Mazzia, V., Khaliq, A. and Chiaberge, M. (2020) Multi-Image Super Resolution of Remotely Sensed Images

- Using Residual Attention Deep Neural Networks. *Remote Sensing*, **12**, Article No. 2207. <https://doi.org/10.3390/rs12142207>
- [55] Song, Q., Xiu, M., Nie, Y., Hu, M. and Liu, C. (2024) CoT-MISR: Marrying Convolution and Transformer for Multi-Image Super-Resolution. *Multimedia Tools and Applications*, **83**, 76891-76903. <https://doi.org/10.1007/s11042-024-18591-4>
- [56] Li, H.A., Zheng, Q.X., Tao, R.L., *et al.* (2023) Review of Image Super-Resolution Based on Deep Learning. *Journal of Graphics*, **44**, 1-15.
- [57] 倪劫, 柳青远, 周莉. 利用改进的 Real-ESRGAN 模型进行历史图像超分辨率重建研究[J]. 信息与管理研究, 2025, 10(1): 65-77.
- [58] Xie, J. and Cheng, X. (2024) Volleyball Game Tactical Data Prediction System Using Computer Vision Technology. *2024 2nd International Conference on Mechatronics, IoT and Industrial Informatics (ICMIII)*, Melbourne, 12-14 June 2024, 588-594. <https://doi.org/10.1109/icmiii62623.2024.00116>
- [59] Nicolson, E., Mohseni, E., Lines, D., Tant, K.M.M., Pierce, G. and MacLeod, C.N. (2024) Towards an In-Process Ultrasonic Phased Array Inspection Method for Narrow-Gap Welds. *NDT & E International*, **144**, Article 103074. <https://doi.org/10.1016/j.ndteint.2024.103074>
- [60] 马钟, 赵歆波, 艾鑫, 张珂. 涡轮叶片 X 射线图像超分辨率重建技术[J]. CT 理论与应用研究(中英文), 2010, 19(1): 41-47.
- [61] Sobek, J., Medina Inojosa, J.R., Medina Inojosa, B.J., *et al.* (2023) MedYOLO: A Medical Image Object Detection Framework. <https://doi.org/10.48550/arXiv.2312.07729>
- [62] Guo, Z., Li, X., Huang, H., Guo, N. and Li, Q. (2019) Deep Learning-Based Image Segmentation on Multimodal Medical Imaging. *IEEE Transactions on Radiation and Plasma Medical Sciences*, **3**, 162-169. <https://doi.org/10.1109/trpms.2018.2890359>
- [63] Avram, O., Durmus, B., Rakocz, N., Corradetti, G., An, U., Nittala, M.G., *et al.* (2024) Accurate Prediction of Disease-Risk Factors from Volumetric Medical Scans by a Deep Vision Model Pre-Trained with 2D Scans. *Nature Biomedical Engineering*, **9**, 507-520. <https://doi.org/10.1038/s41551-024-01257-9>