

A Comprehensive Survey on Super-Resolution Techniques in Medical Imaging

Girija Rani G^{1*} & Prof. S. Jyothi¹

¹ Dept. Of Computer Science, Sri Padmavati Mahila Visvavidyalayam (Women's University), Tirupati, Andhra Pradesh, India

ARTICLE INFO

2026 Volume 3

<https://www.doi.org/ccrcr.2025.tgc.0366>

Article History

Received: June 10, 2024

Accepted: Sep 21, 2024

Published: Dec 15, 2024

Citation: Girija Rani, G., & Jyothi, S. (2024). A comprehensive survey on super-resolution techniques in medical imaging. *Chronicles of Clinical Reviews and Case Reports*, The Geek Chronicles, 3, 1–27.

Copyright: © 2024 Girija Rani G. this is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Keywords: Medical Image Super-Resolution; Deep Learning; Machine Learning; CNN; GAN; Vision Transformer.

ABSTRACT

Medical image super-resolution (SR) has emerged as a critical technique for enhancing diagnostic capabilities while addressing limitations in imaging hardware, scan time, and patient comfort. This survey provides a comprehensive review of machine learning (ML) and deep learning (DL) approaches for medical image SR, covering classical methods, traditional ML techniques, and state-of-the-art deep learning architectures. We systematically analyze SR methods across multiple medical imaging modalities including CT, MRI, ultrasound, microscopy, and nuclear imaging. Through extensive comparison of network architectures, loss functions, and evaluation metrics, we highlight the evolution from sparse coding and dictionary learning to advanced convolutional neural networks (CNNs), generative adversarial networks (GANs), and vision transformers. We discuss critical challenges including data scarcity, domain generalization, clinical validation, and real-world deployment. Finally, we outline promising future directions including self-supervised learning, federated learning, and explainable AI for medical SR.

Introduction

Medical imaging has become an indispensable component of modern clinical practice, enabling clinicians to visualize and assess anatomical structures and physiological processes non-invasively. Imaging modalities such as Magnetic Resonance Imaging (MRI), Computed Tomography (CT), Ultrasound, Positron Emission Tomography (PET), and Digital Pathology are integral to disease diagnosis, treatment planning, and clinical decision-making [1]. The diagnostic accuracy of these modalities is directly influenced by image quality, particularly spatial resolution, which determines the ability to visualize fine anatomical details such as microvasculature, lesions, or tissue boundaries.

Despite continuous improvements in acquisition hardware and reconstruction algorithms, medical images are often acquired at limited resolution. Physical and clinical constraints—including scan time, patient motion, hardware limits, and radiation dose—pose significant barriers to high-resolution imaging [4]. In CT imaging, increasing spatial resolution is often achieved at the cost of higher radiation exposure, which may be clinically unacceptable [3]. In MRI, high-resolution scans require longer acquisition times, resulting in patient discomfort, increased risk of motion artifacts, and lower throughput. Ultrasound imaging is fundamentally limited by transducer bandwidth and signal attenuation.

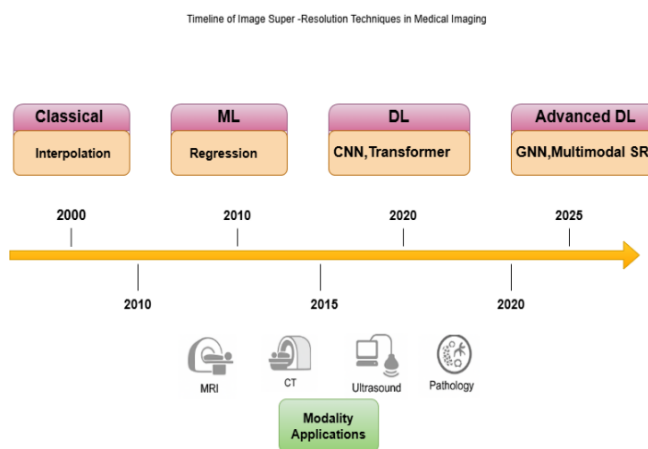


Figure 1. Timeline of medical image super-resolution (SR) techniques, illustrating the progression from classical interpolation and ML regression methods to modern DL and advanced GNN-based multimodal SR approaches across MRI, CT, ultrasound, and pathology modalities.

Numerous studies have demonstrated the clinical benefits of super-resolution (SR). In MRI, SR techniques can enhance tumor visualization and improve volumetric quantification accuracy [7]. In CT imaging, SR can allow for reduced radiation dose while preserving image sharpness [8]. In ultrasound, SR has been applied to improve vascular imaging and microstructure visualization beyond the physical diffraction limit [5]. In digital pathology, SR can help reconstruct gigapixel-level whole-slide images, improving the accuracy of computational diagnosis. These examples highlight how SR can improve both diagnostic confidence and clinical efficiency. The medical imaging field has increasingly adopted super-resolution (SR) techniques to overcome

modality-specific limitations. In MRI, SR has been applied to improve spatial resolution for tumor detection, brain imaging, and functional mapping [5]. In CT, SR can reduce the required radiation dose by reconstructing high-resolution images from low-dose acquisitions [8]. Ultrasound SR enables super-resolved visualization of vascular networks, expanding the utility of non-invasive imaging [6]. In digital pathology, SR enhances image quality for automated analysis tasks such as segmentation and classification.

Despite these advances, several challenges remain. Data scarcity is a persistent bottleneck, as paired low-resolution–high-resolution (LR–HR) datasets are difficult to obtain in clinical settings due to privacy and acquisition constraints [18]. Domain generalization

remains limited, as SR models trained on data from one scanner or institution often perform poorly on others. Standard evaluation metrics such as PSNR and SSIM may not correlate well with clinical utility, necessitating task-specific and perceptual evaluation strategies. Furthermore, the interpretability and reliability of deep learning-based SR methods remain active areas of research.

Main contributions of this survey

Although several reviews on super-resolution (SR) exist, few provide a comprehensive, modality-aware synthesis of classical, machine learning (ML), and deep learning (DL) methods specifically tailored to medical imaging. The main contributions of this survey are:

- A structured review of SR methods from classical interpolation to ML and DL techniques, emphasizing their clinical motivations and methodological evolution.
- A taxonomy of SR approaches categorized by imaging modality (MRI, CT, Ultrasound, Pathology, and others) with emphasis on architecture design, training strategies, and loss functions.
- A comparative analysis of SR performance across modalities, including reconstruction quality, computational cost, and clinical utility.
- A discussion of open research challenges and future opportunities, including self-supervised learning, domain adaptation, and explainable SR models for clinical deployment.

In summary, SR represents a powerful approach to enhancing medical image quality while minimizing acquisition costs and risks. This survey aims to serve as a comprehensive reference for researchers and practitioners at the intersection of medical imaging and artificial intelligence, supporting both methodological innovation and clinical translation.

I. Fundamentals of Image Resolution

A. Problem Definition and Terminology

Image Super-Resolution (SR) refers to the process of reconstructing a high-resolution (HR) image from one or more low-resolution (LR) inputs [19]. The goal is to recover lost high-frequency information, such as edges and fine anatomical details, that are typically degraded during image acquisition. SR is a crucial task in medical imaging, as it enhances the visibility of structures and supports improved diagnosis, image analysis, and computer-aided decision-making [6].

The SR problem is often modeled as an inverse problem. Given an LR image I_{LR} , the objective is to estimate its corresponding HR image I_{HR} such that the difference between I_{HR} and the unknown ground truth is minimized. Typically, an LR image can be mathematically expressed as:

$$I_{LR} = D(H(I_{HR})) + n \quad (1)$$

where H denotes a blur operator, D a downsampling function, and n additive noise [21]. Recovering I_{HR} from I_{LR} is inherently ill-posed since multiple HR images can correspond to the same LR input.

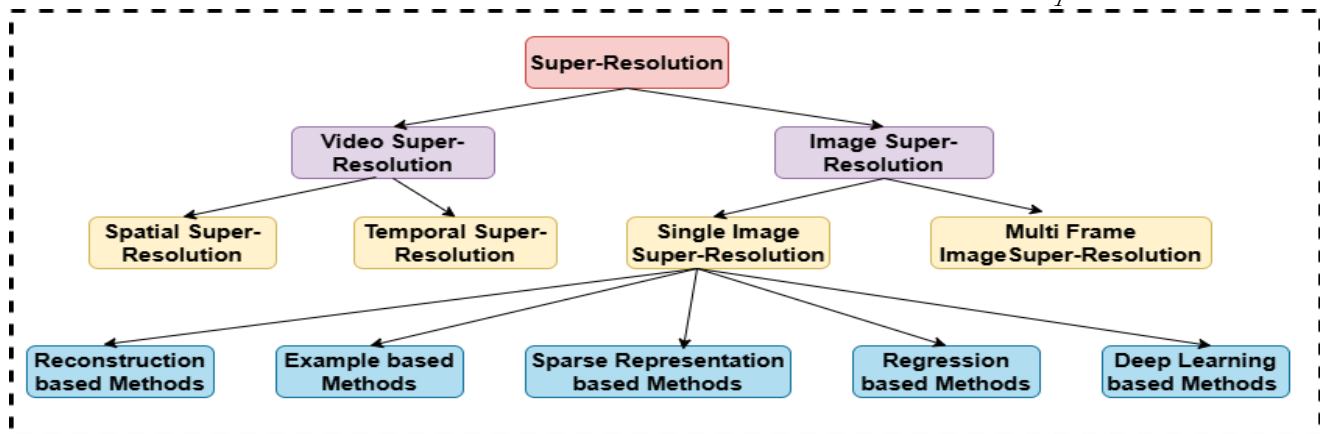


Figure 2. Taxonomy of super-resolution (SR) techniques, categorized into video and image SR. Image SR further includes single-image and multi-frame SR approaches, encompassing reconstruction-, example-, sparse representation-, regression-, and deep learning-based methods.

Terminology.

In the super-resolution (SR) literature, the *upsampling factor* refers to the magnification ratio between low-resolution (LR) and high-resolution (HR) images, commonly $\times 2$, $\times 4$, or $\times 8$ [13]. *Single Image Super-Resolution (SISR)* uses one LR input, whereas *Multi-Frame Super-Resolution (MFSR)* leverages multiple temporally or spatially aligned LR images [22]. *Reference-based SR (RefSR)* integrates external high-quality images as guidance, and *Video SR (VSR)* exploits temporal continuity.

B. Evaluation Metrics

Evaluation of super-resolution (SR) performance is commonly based on Peak Signal-to-Noise Ratio (PSNR) and Structural Similarity Index (SSIM) [27]. However, in medical imaging, these metrics may not accurately reflect clinical utility. Perceptual metrics such as LPIPS and task-based evaluations like segmentation accuracy or radiomic stability are increasingly being used [28].

Notational Conventions.

Let $X \in \mathbb{R}^{h \times w}$ denote a low-resolution (LR) image and $Y \in \mathbb{R}^{h^* \times w^*}$ the corresponding high-resolution (HR) image, with $H = s \times h$ and $W = s \times w$, where s is the scaling factor. The SR mapping function f^θ , parameterized by θ , is learned to minimize:

$$\hat{Y} = f_\theta(X) \quad (2)$$

where \hat{Y} is the reconstructed high-resolution image.

Evaluation metrics play a critical role in measuring the performance of super-resolution (SR) algorithms. They determine not only the fidelity of the reconstructed image but also its perceptual and clinical utility. In medical imaging, choosing appropriate evaluation criteria is especially important because purely numerical improvements may not translate to better diagnostic outcomes [27].

Peak Signal-to-Noise Ratio (PSNR).

PSNR is the most widely used quantitative metric in image super-resolution (SR) tasks. It measures the pixel-wise difference between the reconstructed SR image (\hat{I}) and the reference high-resolution (HR) image (I). It is defined as:

$$\begin{aligned} & \text{PSNR} \\ &= 10 \cdot \log_{10} \left(\frac{\text{MAX}_I^2}{\text{MSE}} \right) \end{aligned} \quad (3)$$

where MAX is the maximum possible pixel value, and MSE denotes the mean squared error. A higher PSNR indicates a lower reconstruction error [27]. Although PSNR is simple and easy to compute, it correlates poorly with perceptual quality in medical imaging scenarios.

Structural Similarity Index (SSIM).

SSIM measures perceptual similarity between two images by considering luminance, contrast, and structural information [27]. It is defined as:

$$\begin{aligned} & \text{SSIM}(I, \hat{I}) \\ &= \frac{(2\mu_I \mu_{\hat{I}} + C_1)(2\sigma_{I\hat{I}} + C_2)}{(\mu_I^2 + \mu_{\hat{I}}^2 + C_1)(\sigma_I^2 + \sigma_{\hat{I}}^2 + C_2)} \end{aligned} \quad (4)$$

where μ and σ represent the mean and standard deviation, and C_1 , C_2 are stability constants. SSIM better captures structural details than PSNR, which is important for detecting small lesions or anatomical features.

Learned Perceptual Image Patch Similarity (LPIPS).

LPIPS evaluates perceptual similarity using deep neural network features [26]. It compares feature activations between I and \hat{I} using a pretrained network such as VGG. LPIPS correlates better with human perception [100] than PSNR or SSIM, making it valuable for evaluating perceptual quality in super-resolution (SR), especially when GAN-based methods are used.

Mean Opinion Score (MOS).

The Mean Opinion Score (MOS) is a subjective evaluation metric that involves radiologists or experts rating image quality on a predefined scale (for example, 1 to 5) [1]. It reflects clinical acceptability more accurately than objective metrics but requires human effort and standardized evaluation protocols. MOS is often used in validation studies of super-resolution (SR) algorithms for CT, MRI, and pathology images.

Perceptual and Fidelity Trade-offs.

Objective fidelity metrics such as PSNR and SSIM may not always align with perceptual or clinical image quality. GAN-based super-resolution (SR) methods often achieve lower PSNR but higher MOS or LPIPS

scores. Therefore, multiple complementary metrics are typically reported to capture different aspects of image quality [28].

Multi-Scale and Region-Based Evaluation.

In medical super-resolution (SR), evaluating specific regions of interest (ROI) such as tumors, vessels, or soft tissue is more meaningful than relying on global averages. Multi-scale SSIM or ROI-based PSNR can provide more clinically relevant insights [7].

Standardized Benchmarks.

To ensure comparability, standardized datasets and evaluation protocols are essential. While natural image super-resolution (SR) has established benchmarks such as Set5, Set14, and DIV2K, medical SR requires modality-specific benchmarks for MRI, CT, ultrasound, and pathology [6].

Visualization and Qualitative Evaluation.

Figures and visual inspections are often used alongside quantitative metrics to highlight improvements in edges, textures, and lesion visibility. High perceptual quality is particularly important in clinical decision-making, even when numerical metrics are not optimal [26].

B. Future Directions in SR Evaluation.

Recent studies have explored integrating both perceptual and clinical utility into the evaluation of super-resolution (SR) methods. Hybrid metrics that combine fidelity, perceptual similarity, and task-based outcomes are being developed to provide a more robust and clinically meaningful assessment of medical SR performance [28].

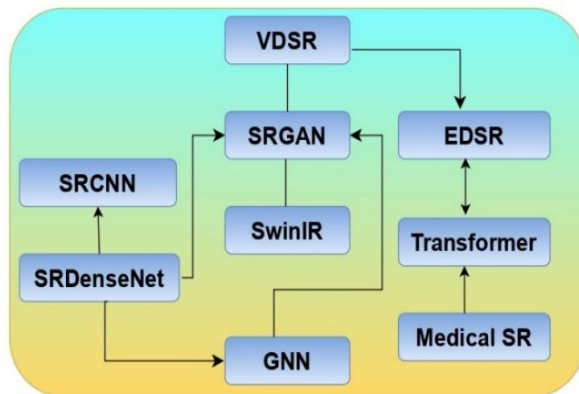


Figure 3. Evolution of deep learning-based super-resolution (SR) models, illustrating the progression from early CNN architectures (SRCNN, SRDenseNet, VDSR) to advanced frameworks (SRGAN, EDSR, SwinIR, Transformer, and GNN), culminating in specialized medical SR applications.

II. Comparison of SR Architectures and Performance

A fair and systematic comparison of super-resolution (SR) architectures is essential for understanding their relative strengths and limitations. In both natural and medical imaging, the performance of SR algorithms is typically evaluated using quantitative metrics such as Peak Signal-to-Noise Ratio (PSNR), Structural Similarity Index (SSIM), and Learned Perceptual Image Patch Similarity (LPIPS), along with qualitative assessments [28].

A. Benchmark Datasets.

Natural image super-resolution (SR) has well-established benchmark datasets such as Set5, Set14, BSD100, and DIV2K. These datasets provide

standardized low-resolution (LR) and high-resolution (HR) pairs, along with defined degradation models, enabling consistent model evaluation [13].

In contrast, medical SR involves modality-specific datasets such as MRI brain images, low-dose CT datasets [95], and high-frequency ultrasound volumes [7] white matter tracts and small lesions. These datasets vary in resolution, noise characteristics, and clinical context, requiring specialized evaluation protocols and careful interpretation of quantitative metrics.

SR in Magnetic Resonance Imaging (MRI)

Magnetic Resonance Imaging (MRI) provides exceptional soft tissue contrast but is often acquired at lower resolution to reduce scan time and mitigate patient discomfort. Super-resolution (SR)

reconstruction enables the enhancement of through-plane resolution, reduction of scan time, and improved depiction of anatomical details such as white matter tracts and small lesions [4].

Deep learning-based SR has been successfully applied to accelerate MRI protocols, producing diagnostic-quality images with significantly reduced acquisition times. Transformer-based SR methods [16] further improve structural fidelity across slices, supporting more accurate segmentation, quantitative analysis, and radiomic feature extraction in clinical workflows.

SR in Computed Tomography (CT)

Computed Tomography (CT) involves the use of ionizing radiation, making the trade-off between resolution and radiation dose clinically critical. Super-resolution (SR) enables the reconstruction [20] of high-resolution (HR) CT images from low-dose acquisitions, thereby reducing patient exposure while maintaining diagnostic quality [3].

GAN-based SR approaches have demonstrated superior performance in enhancing edge sharpness and lesion detectability, particularly in lung nodule screening and cardiac imaging [98]. Furthermore, SR can assist in recovering fine vascular structures and improving image quality for accurate radiotherapy planning.

SR in Ultrasound Imaging

Ultrasound is a real-time and low-cost imaging modality but suffers from low spatial resolution due to transducer and hardware limitations. Super-resolution (SR) techniques enhance ultrasound resolution, enabling improved visualization of microvascular structures, small lesions, and fetal anatomical details [5]. Recent advancements have introduced real-time SR models capable of running on edge devices, making them highly suitable for point-of-care and telemedicine applications. In addition, transformer-based and hybrid network architectures are being explored to achieve robust SR performance under challenging conditions such as speckle noise and motion artifacts.

SR in Digital Pathology and Microscopy

Digital pathology and microscopy produce gigapixel-scale images; however, high-resolution imaging is often expensive and time-consuming. Super-resolution (SR) enables the upscaling of lower-resolution scans while preserving fine cellular and subcellular details, thereby

allowing faster slide scanning and efficient computational analysis [34].

Deep learning-based SR methods have been successfully applied in hematology, oncology, and histopathology workflows, supporting AI-driven diagnosis, biomarker quantification, and automated grading systems. These advancements enhance both the speed and accuracy of digital pathology pipelines in clinical and research environments [35].

SR for Functional and Dynamic Imaging

Super-resolution (SR) is increasingly being applied in functional MRI (fMRI), dynamic contrast-enhanced MRI, and 4D CT to address the trade-off between spatial and temporal resolution. By enhancing spatial resolution post-acquisition, SR enables shorter scan times while preserving essential functional and physiological information. This capability is particularly valuable in cardiac imaging, perfusion studies, and neuroimaging research, where temporal precision and anatomical detail are both crucial.

SR for Segmentation, Detection, and Radiomics

Super-resolution (SR) serves as a powerful preprocessing step to enhance downstream artificial intelligence (AI) tasks such as segmentation, registration, and classification. Higher-resolution images improve boundary delineation, particularly in tumor segmentation, vessel analysis, and organ contouring. In addition, SR contributes to more stable and reproducible radiomic feature extraction, which is highly sensitive to image resolution and voxel size [8].

Edge Deployment and Real-Time Applications

Lightweight convolutional neural network (CNN) and efficient Transformer models have made it possible to perform super-resolution (SR) in real-time clinical environments. These advancements enable SR on portable ultrasound devices, intraoperative surgical guidance systems, and low-resource medical setups. Such real-time implementations improve accessibility and accelerate clinical decision-making, particularly in emergency care and rural healthcare settings where computational resources are limited.

B. Applications in Medical Imaging.

Sparse coding methods were widely applied in Magnetic Resonance Imaging (MRI), Computed Tomography (CT), and microscopy-based super-resolution (SR) before convolutional neural networks (CNNs) [85][86]

became dominant in the field. These methods improved through-plane resolution, edge sharpness, and noise suppression in reconstructed images while remaining computationally efficient compared to early deep learning models [41].

Limitations

Despite their strong performance, sparse coding methods rely heavily on handcrafted features, are sensitive to the quality of the learned dictionary, and often lack generalization to unseen domains. These limitations paved the way for deep learning-based approaches that automatically learn hierarchical feature representations directly from data.

C. Manifold Learning & Patch-Based ML Methods

Manifold learning-based super-resolution (SR) approaches assume that low-resolution (LR) and high-resolution (HR) image patches lie on low-dimensional manifolds embedded within high-dimensional spaces [47][42]. The key concept is to learn the intrinsic structure of these manifolds and utilize it to reconstruct HR patches from their LR counterparts. Unlike classical interpolation or linear regression approaches, manifold learning methods better preserve structural relationships, making them particularly effective for medical images that contain complex anatomical patterns and textures.

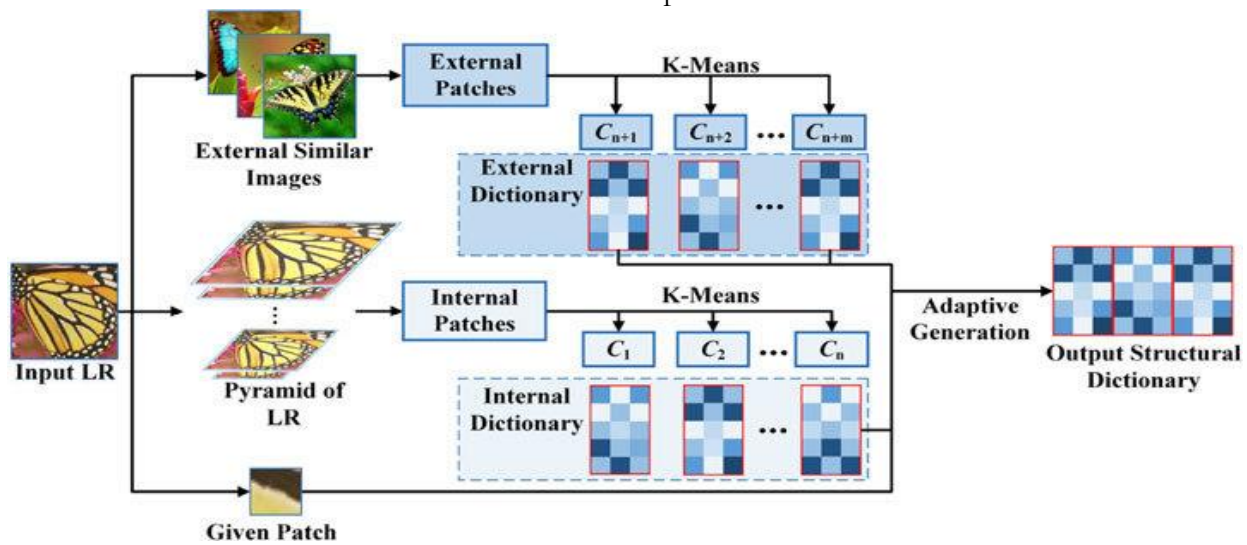


Figure 4. Illustration of an adaptive structural dictionary learning framework, where both internal and external patches are extracted and clustered using K-means to form internal and external dictionaries. These are then adaptively fused to generate an optimized output structural dictionary for image super-resolution.

Motivation

Medical images often contain highly structured patterns such as vessels, organ boundaries, and tissue textures that can be more effectively represented on nonlinear manifolds. Mapping low-resolution (LR) patches to high-resolution (HR) manifolds enables super-resolution (SR) algorithms to leverage geometric relationships rather than relying solely on pixel intensity similarity [43].

Locally Linear Embedding (LLE) for SR.

Chang et al. [43] introduced manifold learning into super-resolution (SR) through the concept of locally linear embedding (LLE). Given a low-resolution (LR) image patch, its k nearest neighbors is identified within

the training dataset, and reconstruction weights are computed accordingly. The same weights are then applied in the high-resolution (HR) manifold to reconstruct the HR patch, as defined by the following equation:

$$y = \sum_{i=1}^k w_i y_i \quad (5)$$

Laplacian Eigenmaps

Belkin and Niyogi [48] proposed the Laplacian Eigenmaps method for manifold embedding, which preserves neighborhood relationships within a low-dimensional space. When applied to super-resolution

(SR), this approach improves structural consistency between low-resolution (LR) and high-resolution (HR) image patches, resulting in better preservation of anatomical features and fine textures.

Manifold Alignment and Mapping.

More advanced super-resolution (SR) methods learn explicit mappings between low-resolution (LR) and high-resolution (HR) manifolds by aligning their geometric structures. This alignment can be achieved using techniques such as kernel regression, manifold alignment, or graph embedding approaches [47]. These methods enhance the ability of SR models to preserve structural coherence and spatial relationships across image scales.

Patch-Based Manifold SR

Patch-based ML models combine manifold learning with patch clustering or anchor-based selection to improve efficiency. In this framework, LR patches are embedded in a low-dimensional manifold, neighbors are selected, and corresponding HR patches are reconstructed through weighted combinations or learned mappings.

Manifold Learning in Medical SR.

Manifold learning methods have been applied to medical super-resolution (SR) to enhance Magnetic Resonance Imaging (MRI) and Computed Tomography (CT) scans, particularly for improving through-plane resolution and preserving anatomical structure. Because these methods model local geometric relationships, they can handle anatomical variability

more effectively than traditional interpolation-based approaches [51][44].

Limitations.

Manifold-based super-resolution (SR) methods require large and diverse training datasets, are sensitive to neighbor selection, and often struggle to generalize to unseen anatomical patterns. In addition, these methods involve computationally expensive nearest-neighbor searches, which limit their scalability. These challenges ultimately led to the adoption of deep learning approaches that can learn complex nonlinear mappings directly from data.

D. Hybrid ML + Model-Based Reconstruction

Hybrid super-resolution (SR) methods combine model-based image reconstruction with machine learning

(ML) priors to achieve improved image quality, better generalization, and enhanced stability [53][45]. These approaches integrate explicit imaging models—such as blur, noise, and down sampling functions—with learned patch-based or manifold-based priors. By doing so, they effectively bridge the gap between traditional inverse problem formulations and modern data-driven learning frameworks.

Motivation.

Pure model-based reconstruction methods rely on analytical priors such as total variation (TV) or edge-preserving regularization, which can over smooth important anatomical structures. On the other hand, purely machine learning-based methods may hallucinate or generate non-physical details that compromise clinical reliability. Hybrid approaches combine the strengths of both paradigms, ensuring physical consistency while enhancing perceptual and diagnostic image quality [11][46].

General Formulation.

Hybrid super-resolution (SR) methods are typically formulated as an optimization problem:

$$\hat{y} = \arg \min_y \|x - H \cdot y\|^2 + \lambda R(y) \quad (6)$$

where x denotes the low-resolution (LR) image, H represents the imaging degradation model (such as blur and down sampling), and $R(y)$ is a regularization term that incorporates learned priors—such as sparse coding, manifold structure, or regression-based constraints.

Patch-Based Learning Priors.

One hybrid approach integrates sparse coding priors into the reconstruction process. For instance, Elad and Aharon [40] incorporated dictionary learning into total variation (TV)-regularized reconstruction, enabling the preservation of fine structural details while maintaining strong data fidelity constraints.

Manifold and Regression Priors.

Protter et al. [53] proposed the use of neighbor embedding priors within a reconstruction framework for super-resolution (SR). In this approach, regression functions learned offline are combined with a forward imaging model to ensure that the reconstructed high-resolution (HR) output, when degraded, remains consistent with the original low-resolution (LR) input. Regularization Techniques.

Hybrid super-resolution (SR) methods often incorporate total variation (TV) regularization, sparsity constraints, or edge-preserving priors alongside learned data-driven priors. This combination is particularly effective in medical imaging [99] applications, where it is crucial to preserve both smoothness and anatomical structures while enhancing fine details [22].

Applications in Medical Imaging.

In Magnetic Resonance Imaging (MRI), hybrid reconstruction techniques improve through-plane resolution by embedding machine learning (ML) priors into iterative reconstruction algorithms. In Computed Tomography (CT), these approaches are used to recover sharp structural details while effectively controlling noise [23][50]. Such methods are particularly attractive in scenarios where large training datasets are unavailable, as the model-based components contribute to improved stability and robustness.

Advantages and Limitations.

Hybrid super-resolution (SR) methods offer superior fidelity, interpretability, and stability compared to purely learning-based approaches. However, their iterative reconstruction procedures can be computationally expensive and are often sensitive to hyperparameter settings. These limitations inspired the development of later deep learning models, such as model-based deep unfolding networks [87][88], which build upon the foundational principles of hybrid SR.

Limitations of ML Methods in Medical Imaging

While classical machine learning (ML) methods for super-resolution (SR)—such as dictionary learning, regression-based mapping, and manifold learning—laid the groundwork for modern deep learning approaches, they exhibit several key limitations when applied to medical imaging [19]. These limitations arise from both methodological constraints and the inherent complexities of clinical imaging domains.

1. Limited Representation Capacity.

ML methods typically rely on handcrafted features, shallow regression models, or linear mapping functions. These are often insufficient to capture the rich anatomical variability present in medical images, particularly for complex modalities such as MRI and CT [10].

2. Domain Shift and Poor Generalization.

Medical data are heterogeneous across scanners, hospitals, and acquisition protocols. Models trained on one dataset often fail to generalize to others due to domain shift. Traditional ML models lack domain adaptation and transfer learning capabilities [22].

3. Dependence on Paired Data and Patch Dictionaries.

Dictionary learning and regression-based SR approaches require high-quality paired low-resolution (LR) and high-resolution (HR) datasets or large patch dictionaries. This is particularly challenging in medical imaging, where true HR ground truth data are not always available [12][49][85].

4. Lack of End-to-End Optimization.

Most ML-based SR methods treat feature extraction, mapping, and reconstruction as separate modules. This modularity often leads to suboptimal results compared to deep learning methods that learn hierarchical features through end-to-end training.

5. Computational Burden at Inference.

Patch search, sparse coding, and neighbor embedding techniques require computationally expensive nearest-neighbor searches and optimization during inference [43]. This significantly limits their clinical applicability, especially in time-sensitive workflows such as intraoperative imaging.

6. Sensitivity to Noise and Artifacts.

ML-based SR methods often amplify acquisition noise or fail to handle motion artifacts due to their reliance on fixed linear mappings or local patch similarity [37][52].

7. Lack of Interpretability and Standardization.

Although ML methods are simpler than deep learning approaches, many still lack formal clinical validation and standardized reconstruction quality assurance protocols. This limits their acceptance in regulated clinical environments [36].

8. Limited Integration with Imaging Physics.

Most traditional ML methods ignore the forward imaging model, resulting in reconstructions that are not physically consistent with scanner acquisition models. This limitation is particularly problematic in CT imaging, where dose constraints and artifact suppression are critical [54][38].

9. Scalability Issues.

As image sizes and modalities increase (for example, 3D MRI and 4D fMRI), the scalability of patch-based ML methods becomes a bottleneck. Storage requirements and computation time grow rapidly with dictionary size and neighbor searches.

10. Lack of Regulatory and Clinical Adoption.

Due to the aforementioned limitations, very few ML-based SR methods have achieved clinical translation or regulatory approval. Most remain confined to research or proof-of-concept studies rather than routine diagnostic workflows.

III. Deep Learning Foundations

Deep learning (DL) has revolutionized image super-resolution (SR) by enabling the learning of complex nonlinear mappings from low-resolution (LR) to high-resolution (HR) images directly from data [14][39]. Unlike classical machine learning (ML) methods that depend on handcrafted features or shallow regression mappings, DL architectures can automatically learn hierarchical feature representations through multi-layer neural networks. This data-driven capability allows DL-based SR models to capture intricate spatial dependencies and recover fine anatomical details that are often lost in traditional reconstruction or interpolation techniques.

This section provides a concise background on the core building blocks used in modern SR architectures, focusing on neural network components, activation functions, and up sampling mechanisms.

Network Components

Most deep learning-based SR networks are composed of fundamental components such as convolutional layers, activation functions, normalization layers, and up sampling operators. These components form the backbone of SR architectures like SRCNN, VDSR, EDSR, SRGAN, and SwinIR.

1. Convolutional Layers.

Convolutional layers form the core building blocks of super-resolution (SR) networks. They extract local image features using learnable filters and shared weights, which enable translation-invariant feature extraction across the image. When multiple convolutional layers are stacked, the network progressively learns more abstract and complex feature

representations, improving its ability to reconstruct fine details and textures [55].

2. Nonlinear Activation Functions.

Nonlinear activation functions such as ReLU, PReLU, and Leaky ReLU introduce nonlinearity into neural networks, allowing them to learn and approximate complex mappings between low-resolution (LR) and high-resolution (HR) representations. Among these, the Rectified Linear Unit (ReLU) is the most widely used due to its simplicity, effectiveness in mitigating vanishing gradients, and computational efficiency [57].

3. Normalization Layers

Batch Normalization (BN) and Instance Normalization (IN) are commonly used to stabilize network training and accelerate convergence by normalizing intermediate feature maps. However, some super-resolution (SR) networks—such as the Enhanced Deep Super-Resolution (EDSR) model—intentionally omit normalization layers to better preserve fine image details and avoid over-smoothing [59].

4. Residual Connections

Residual connections enhance gradient flow and allow for the training of deeper neural network architectures without suffering from vanishing gradient problems. They are a fundamental component of modern super-resolution (SR) networks such as the Very Deep Super-Resolution (VDSR) and Enhanced Deep Super-Resolution (EDSR) models, enabling stable optimization and improved reconstruction performance [30].

5. Upsampling Layers.

Upsampling plays a critical role in super-resolution (SR) by reconstructing high-resolution (HR) images from low-resolution (LR) feature representations. Common upsampling strategies include:

- Deconvolution / Transposed Convolution: A learnable upsampling technique that increases spatial resolution through trainable filters.
- Sub-pixel Convolution (PixelShuffle): Rearranges feature maps to efficiently increase image resolution with reduced computational cost [31][58].
- Interpolation + Convolution: Combines classical interpolation methods (e.g., bicubic or bilinear) [9] with subsequent convolutional refinement to enhance visual fidelity.

6. Skip Connections and Dense Blocks.

Skip connections transmit low-level feature information to deeper layers, helping preserve fine image details and improving gradient flow during training. Dense blocks, as introduced in architectures such as DenseNet and SRDenseNet, further enhance feature reuse and encourage efficient information propagation across layers, resulting in better reconstruction quality and network stability [32][60].

7. Attention Modules.

Channel and spatial attention modules enable the network to focus selectively on the most informative regions and feature channels, thereby improving perceptual quality and reconstruction accuracy. These

mechanisms help super-resolution (SR) models emphasize clinically relevant image areas while reducing redundancy in feature learning [26][61].

8. Loss Functions.

Although discussed in detail in later sections, it is important to note that deep learning-based super-resolution (SR) models typically employ a combination of loss functions to guide reconstruction quality. These include pixel-wise losses such as Mean Squared Error (MSE), perceptual losses that capture visual fidelity, adversarial losses used in Generative Adversarial Networks (GANs), and task-specific losses designed for clinical objectives such as segmentation or detection accuracy.

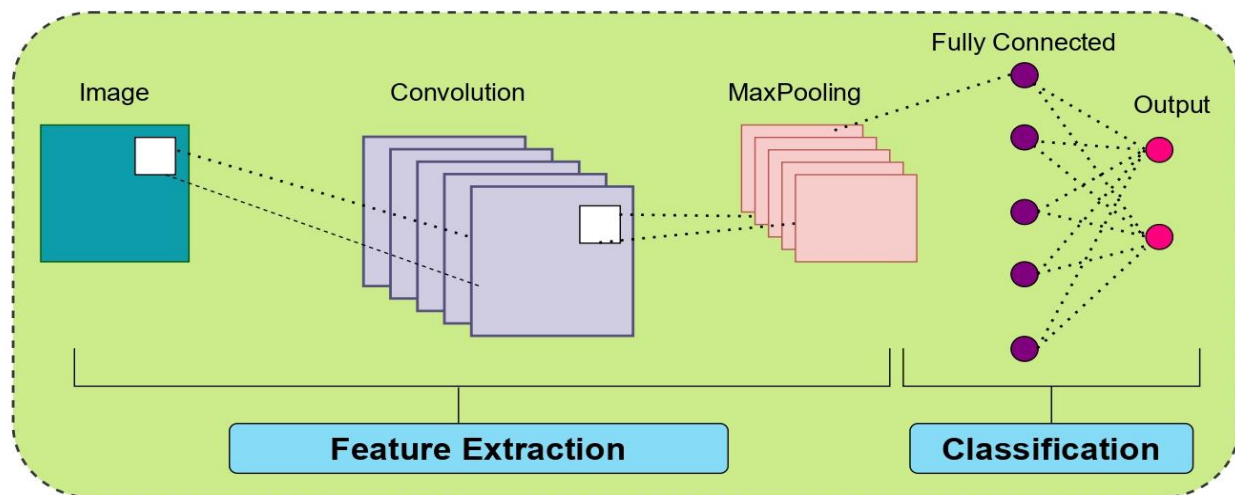


Figure 5. General architecture of a convolutional neural network (CNN), illustrating the feature extraction phase using convolution and max-pooling layers, followed by a fully connected network for classification and output prediction.

Convolutional Neural Networks (CNNs)

CNNs form the backbone of most deep learning-based super-resolution (SR) architectures. They are highly effective in modeling local spatial dependencies and learning hierarchical feature representations from image data, making them particularly suitable for reconstructing fine anatomical structures in medical imaging [55][62].

Core Principle.

Convolutional Neural Network (CNN)-based super-resolution (SR) models learn a nonlinear mapping function between low-resolution (LR) and high-resolution (HR) images. This can be mathematically expressed as:

$$\hat{y} = f\theta(x) \quad (7)$$

where x denotes the LR input, \hat{y} represents the reconstructed HR image, and θ corresponds to the learnable network parameters. Stacked convolutional layers progressively extract hierarchical features from the input, enabling the network to recover fine details and enhance overall image fidelity.

Representative CNN SR Models

SRCNN [56]: The first CNN-based super-resolution (SR) model, consisting of a three-layer network that performs patch extraction, nonlinear mapping, and image reconstruction.

VDSR [30]: A very deep SR model that incorporates residual learning to enable stable optimization and faster convergence.

EDSR [15]: An enhanced deep SR network that removes batch normalization layers to produce sharper and more detailed images.

SRDenseNet [101]: Utilizes densely connected convolutional layers to encourage feature reuse and improve gradient propagation throughout the network.

Generative Adversarial Networks (GANs)

GANs have greatly improved the perceptual quality of super-resolution (SR) reconstructions by learning to generate realistic textures and fine details. Through the adversarial learning framework, GAN-based SR models encourage the generator to produce high-resolution images that are perceptually indistinguishable from real images [64][69].

GAN Structure

A Generative Adversarial Network (GAN) consists of two primary components: a generator (G) that reconstructs high-resolution (HR) images from low-resolution (LR) inputs, and a discriminator (D) that distinguishes real HR images from generated ones. The training objective is formulated as:

$$\min_G \max_D \mathcal{L}_{GAN(G,D)} = E_y[\log D(y)] + E_x[\log(1 - D(G(x)))] \quad (8)$$

In this framework, the generator aims to produce realistic HR images that can “fool” the discriminator, while the discriminator learns to differentiate between genuine and generated samples. This adversarial training process drives the generator toward producing perceptually convincing images.

SRGAN [14] introduced adversarial training for super-resolution (SR), significantly improving perceptual sharpness and visual realism in reconstructed images.

Variants.

ESRGAN [29]: Incorporates enhanced perceptual loss and a relativistic discriminator to produce finer textures and improved detail consistency.

pGAN [65]: Applied to Magnetic Resonance Imaging (MRI) super-resolution, leveraging adversarial learning for high-quality image reconstruction.

CycleGAN: Enables unpaired SR learning by translating between low-resolution and high-resolution domains without requiring paired datasets.

Advantages.

Generative Adversarial Networks (GANs) are capable of producing sharp, realistic, and perceptually convincing images. They are widely adopted in medical super-resolution (SR) applications such as Magnetic Resonance Imaging (MRI) reconstruction, Computed Tomography (CT) [89] denoising, and digital pathology image enhancement, where high visual fidelity is crucial for accurate diagnosis.

Limitations.

Despite their effectiveness, GANs are challenging to train due to adversarial instability and sensitivity to hyperparameter selection. They may also hallucinate non-existent structures, which can compromise clinical reliability. Achieving a balance between perceptual sharpness and anatomical fidelity requires careful tuning of loss functions and training strategies.

Transformers and GNN/GCN

Recently, Transformer and Graph Neural Network (GNN) [17] models have demonstrated exceptional performance in super-resolution (SR) by capturing long-range dependencies and leveraging global contextual information [67].

Transformers in SR.

Transformers use self-attention mechanisms to model relationships between distant pixels, allowing them to reconstruct fine textures and maintain global image consistency. Architectures such as SwinIR [33] and IPT (Image Processing Transformer) [68] have achieved state-of-the-art performance in SR tasks, outperforming many convolution-based networks.

Advantages.

- Capture global contextual dependencies
- Highly flexible and scalable architecture
- Perform well on large and high-dimensional medical images

GNN/GCN in SR.

Graph Neural Network (GNN)-based SR methods represent image patches or anatomical regions as graph nodes. Graph convolution operations propagate information across these nodes, enabling efficient modeling of complex spatial relationships and irregular structures often found in medical images [80].

Applications.

- Transformer-based SR for 3D MRI and CT reconstruction

- Graph Convolutional Network (GCN)-based SR for anatomical structure preservation, such as vessel or cortical surface reconstruction

Loss Functions Overview

Loss functions play a critical role in guiding super-resolution (SR) networks toward producing clinically meaningful reconstructions. In medical SR, a combination of pixel-wise, perceptual, and adversarial losses is often used to balance fidelity and visual realism [14][71].

1. Pixel-Wise Loss.

Pixel-wise loss functions, such as Mean Squared Error (MSE) and L1 loss, are among the most commonly used in super-resolution (SR) tasks. They measure the average difference between the reconstructed high-resolution (HR) image and the ground truth image, ensuring pixel-level fidelity.

$$L_{MSE} = \frac{1}{N} \sum_{i=1}^N \|\mathbf{y}_i - \hat{\mathbf{y}}_i\|^2 \quad (9)$$

where \mathbf{y}_i represents the ground truth HR pixel value, $\hat{\mathbf{y}}_i$ is the corresponding reconstructed pixel value, and \mathbf{N} is the total number of pixels. MSE emphasizes overall image fidelity but may produce overly smooth results, while L1 loss tends to preserve sharper edges.

2. Perceptual Loss.

Perceptual loss is computed based on deep feature representations extracted from a pre-trained network (such as VGG). It encourages perceptual similarity between the reconstructed and ground truth images, aligning them in feature space rather than relying solely on pixel-level accuracy [77][72].

3. Adversarial Loss.

Adversarial loss is used in Generative Adversarial Network (GAN)-based super-resolution (SR) models

to enhance realism and texture quality. It helps the generator produce images that are perceptually closer to real high-resolution images [14].

4. Edge / Gradient Loss.

Edge or gradient-based loss functions focus on preserving structural and textural details in images. They are particularly valuable in medical imaging for maintaining vessel boundaries, lesion edges, and other fine anatomical structures.

5. Task-Based Loss.

In medical SR applications, task-based loss functions are linked to specific downstream clinical objectives such as segmentation accuracy or disease classification performance [66]. This ensures that the reconstructed images are not only visually realistic but also clinically useful.

6. Combined Objectives.

Modern SR networks typically use a combination of multiple loss functions to balance fidelity, perceptual quality, and realism. The total loss can be expressed as:

$$L_{total} = \lambda^1 L_{pixel} + \lambda^2 L_{perc} + \lambda^3 L_{adv} + \dots \quad (10)$$

IV. ML and DL Architectures in Medical Super-Resolution (SR)

Machine learning (ML) and deep learning (DL) architectures form the foundation of super-resolution (SR) methods in medical imaging. While ML-based methods rely on handcrafted priors, patch similarity, and sparse representation, DL approaches learn hierarchical features end-to-end, enabling superior reconstruction quality. Hybrid methods combining both have recently emerged to leverage the advantages of each [74].

Table 1. Chronological Literature Review of Machine Learning–Based Medical Image Super-Resolution (2010–2019).

No.	Author(s)	Year	Model / Method	Dataset/ Modality	Evaluation Metrics	Remarks / Future Scope
1	Jianchao Yang et al.	2010	Sparse Representation SR (coupled LR/HR dictionaries)	Natural images (Parthenon, Lena, Statue)	RMSE	Foundation for dictionary-learning SR; adapted later for MRI/CT
2	Van Reeth et al.	2012	MRI SR Survey (TV, Bayesian, POCS)	MRI phantoms and clinical datasets	PSNR, SSIM, SNR	Highlighted motion-free and anisotropic SR

3	Fang et al.	2013	SBSDI — Sparse Denoising + Interpolation for OCT	Retinal SDOCT (AREDS2)	PSNR, CNR	Introduced sparse priors for OCT SR; multimodal potential
4	Wang et al.	2014	Sparse Representation Reconstruction (SRR) for MRI	MRI sequences (Skull Base, Aorta, Knee)	PSNR	Proposed 3D SR with SSIM and clinical validation
5	Thapa et al.	2014	SR Algorithm Comparison for Retinal SR	Fundus images (Nidek AFC-230)	PSNR, SSIM	Suggested blind deconvolution and task-specific priors
6	Zhang et al.	2015	SRNINL — Sparse + Nonlocal MR SR	BrainWeb, ADNI	PSNR, SSIM	Nonlocal prior integration; improved computational efficiency
7	Tang et al.	2015	Sparse SR for Mass Spectrometry Imaging (MSI)	Rat brain (AFMI-MSI)	PSNR, RMSE	Introduced MSI SR; proposed 3D tissue-specific dictionaries
8	Wang et al.	2015	Dual Dictionary Learning (DDL) for Medical SR	CT, MRI datasets	PSNR, SSIM	Encouraged multimodal and physics-aware SR
9	Indu D. et al.	2015	Sparse SR for Fundus Fluorescein Angiography (FFA)	FFA clinical images (Amrita Institute)	Entropy, DIV	Used perceptual metrics; suggested GPU acceleration
10	Huynh et al.	2016	Structured RF + Auto-context for MR→CT SR	ADNI MR-CT paired scans	PSNR, MAE	Bridged MRI-CT SR; motivated hybrid RF-DL methods
11	Jog et al.	2017	REPLICA — Regression Ensembles for Contrast Agreement	Multi-contrast MRI (T1w, T2w, FLAIR)	PSNR, SSIM	Developed ensemble regression; potential hybrid CNN use
12	Zheng et al.	2017	Multi-contrast MRI SR (NEDI extension)	BrainWeb, real MRI	PSNR, RLNE	Extended NEDI with GPU optimization
13	Fang et al.	2017	SSR — Segmentation-based Sparse SR for OCT	Retinal SDOCT (41 subjects)	PSNR, CNR	Added shape-adaptive patches; cross-modality SR possible
14	Jiang et al.	2018	CT SR via Sparse Coding + Global Reconstruction	Dental CBCT, XCAT phantom	PSNR, SSIM	Validated feasibility of 3D SR in CBCT imaging
15	Wei et al.	2018	Multi-dictionary Sparse Coding + RF SR	MRI, CT (Siemens Gallery)	PSNR	Hybrid sparse + RF; proposed 4D SR
16	Abbasi et al.	2018	NWSR — Nonlocal Weighted Sparse SR for OCT	Retinal OCT (28 subjects)	PSNR, CNR, ENL	Improved denoising-integrated SR; GPU parallelization
17	Jiang et al.	2018	FFA SR Benchmark (10 SISR Algorithms)	Canon CF-60DSi Fundus images	PSNR, SSIM	Released FFA SR benchmark dataset
18	Dou et al.	2018	Random Forest Regression SR	MRI/CT datasets	PSNR, SSIM, runtime	Lightweight regression-based SR baseline
19	Zhang et al.	2019	MDL + NLSS SR	MRI, CT	PSNR, SSIM	Integrated deep priors with local self-similarity

20	Hu et al.	2019	PET SR via Coupled Dictionary Learning + RF	Whole-body PET lymphoma data	PSNR, SSIM, RMSE	3D/4D PET SR; motion compensation with DL integration
----	-----------	------	---	------------------------------	------------------	---

Table 2. Chronological Literature Review of Deep Learning–Based Medical Image Super-Resolution Methods (2016–2025).

No.	Author(s)	Year	Model / Method	Dataset/Modality	Evaluation Metrics	Remarks / Future Scope
1	Jiwon Kim et al.	2016	VDSR (Very Deep Super-Resolution)	Set5, Set14, Urban100 (natural images)	PSNR, SSIM	Foundation for CNN-based SR; extendable to medical domains
2	Akshay S. Chaudhari et al.	2018	DeepResolve	Knee MRI (DESS)	PSNR, SSIM, MOS	Transferable to other MRI contrasts and anatomies
3	Honggang Chen et al.	2018	CISRDCNN	Set5, Set14	PSNR, SSIM	Extend to compression and medical codec SR
4	Shengxiang Zhang et al.	2019	FMISR	Retinal DRIVE, Abdomen, Brain, Knee	PSNR, Runtime	Introduce GAN/perceptual losses; expand to 3D
5	Qing Lyu et al.	2019	GAN-CPCE / GAN-CIRCLE	IXI MRI	PSNR, SSIM	Scale to 3D GAN SR with improved realism
6	Chi-Hieu Pham et al.	2019	3D CNN SR Network	Kirby 21, NAMIC MRI	PSNR, SSIM	Investigate perceptual SR and registration-free MRI SR
7	Tae-Hyung Kim et al.	2020	SRDenseNet for SISR	CDT Phantom (CT)	FWHM, CNR	Realistic CT validation with better spatial fidelity
8	Yuchong Gu et al.	2020	MedSRGAN	LUNA16 CT, Brain MRI	PSNR, SSIM, MOS	Extend to multimodal and Transformer-based SR
9	Yue Yu et al.	2021	WFSAN	Chest datasets X-ray	PSNR, SSIM	Integrate wavelet and frequency decomposition for SR
10	Zikang Wei et al.	2021	Improved SRGAN	DIV2K, Remote Sensing (RS) images	PSNR, SSIM	Lightweight design and multimodal adaptation
11	Heng-Sheng Chao et al.	2022	SRCNN/VDSR/EDSR for thin-slice SR	Clinical Chest CT	PSNR, SSIM	Extendable to other imaging modalities
12	Kejie Lyu et al.	2022	JSENet	MIT-Adobe FiveK dataset	PSNR, SSIM, LPIPS	Combine CNN and Transformer hybridization
13	Rui Cheng et al.	2022	FDSR	DIV2K, Set5	PSNR, SSIM	Integrate domain-specific priors for medical SR
14	Waqar Ahmad et al.	2022	Progressive Multi-Scale GAN	Retinal, MRI, Ultrasound	PSNR, SSIM	Develop lightweight 3D SR architectures
15	Jay Shah et al.	2024	LDM-RR	ADNI, OASIS-3 (PET)	RC, SUVR	Optimize diffusion-based SR inference for PET imaging

16	Sevara Mardieva et al.	2024	DRFDN	DIV2K, Set5	PSNR, FLOPs	Adapt to medical datasets and lower FLOPs
17	Alireza Esmailzahi et al.	2025	CLBSR	DIV2K, RealSR	PSNR, SSIM	Lightweight Transformer for Blind SR
18	Qiang Zhang et al.	2025	LRTENet	Hyperspectral datasets	PSNR, SSIM, SAM	Integrate global Transformer modules
19	Ke Xu et al.	2025	SCAN	DIV2K, Set5	PSNR, SSIM	Edge deployment and knowledge distillation
20	Masuma Aktar et al.	2025	FA-PWSR	DIV2K, Set5	PSNR, SSIM, LPIPS	Perceptual Transformer-based lightweight SR
21	Khushboo Singla et al.	2025	D-RRDN	CAMUS, LFW	PSNR, SSIM	Incorporate Transformer-based attention
22	Feiwei Qin et al.	2025	InfraFFN	IR700, FLIR (Infrared)	PSNR, SSIM	Lightweight FFN adaptation for IR SR
23	Garas Gendy et al.	2025	GMN	DIV2K, Set5	PSNR, SSIM	Quantization and pruning for fast inference
24	Varsha Singh et al.	2025	MSARCNN	DIV2K, RealSR	PSNR, SSIM	Extend multi-scale CNN SR to medical domain
25	Shizhuang Weng et al.	2025	DWTN + AMResNet	Hyperspectral wheat data	PSNR, SSIM, ACC	Fusion of segmentation and RefSR methods

Table 3. Chronological Literature Review of Deep Learning, GAN, and Transformer-Based Medical Image Super-Resolution Methods (2017–2025).

No.	Author(s)	Year	Model / Method	Dataset/Modality	Evaluation Metrics	Remarks / Future Scope
1	Kenji Suzuki	2017	Overview of DL in Medical Imaging	CT, X-ray, MRI, PET	AUC, Sensitivity, Specificity, Dice	Lightweight DL, multimodal integration, clinical deployment
2	June-Goo Lee et al.	2017	DL in Medical Imaging: Review	CT, MRI, X-ray, PET	AUC, Dice, Accuracy	GANs, Transformers, self-supervised multimodal fusion
3	Dinggang Shen et al.	2017	Review of DL in Medical Image Analysis	MRI, CT, PET, Ultrasound	Dice, AUC, SSIM	Transformers, explainable AI, federated clinical DL
4	Yuhua Chen et al.	2018	DCSRN (3D Densely Connected SR Network)	HCP 3D MRI (T1-w)	PSNR, SSIM, NRMSE	Adversarial loss; hybrid Transformer SR
5	Justin Ker et al.	2018	DL Applications in Medical Imaging	CT, MRI, PET, X-ray, US	AUC, Dice, Accuracy	GANs/VAEs, lightweight transfer learning

6	Shengxiang Zhang et al.	2019	FMISR (Fast Medical Image SR)	Retinal DRIVE, Brain MRI, CT	PSNR, Runtime	GAN/Perceptual loss, 3D Transformer SR
7	Dwarikanath Mahapatra et al.	2019	P-SRGAN (Progressive GAN + Triplet Loss)	EYEPACS, Sunnybrook Cardiac MRI	PSNR, SSIM, AUC	Self-supervised, perceptual Transformer SR
8	Wenming Yang et al.	2019	DL for SISR: Review	DIV2K, Set5, Set14, Urban100	PSNR, SSIM, LPIPS	Lightweight CNNs, realistic degradation modeling
9	Mingyu Kim et al.	2019	Deep Learning in Medical Imaging	X-ray, CT, MRI, PET	AUC, Sensitivity, Specificity	Unsupervised DL, Transformer fusion
10	Mohammad H. Hesamian et al.	2019	CNN, FCN, U-Net, V-Net	LUNA16, SLiver07, ISBI-2016	Dice, IoU, Accuracy	Hybrid CNN–Transformer, unified benchmarks
11	Grant Haskins et al.	2020	DL for Image Registration: Survey	MRI, CT, US, X-ray, PET	TRE, DSC, MI, PSNR	Transformer/GAN for motion-aware registration
12	Chenyu You et al.	2020	GAN-CIRCLE	Tibia Micro-CT, Mayo Clinic CT	PSNR, SSIM, MOS	GAN + Transformer hybrid; adversarial robustness
13	Yabo Fu et al.	2020	DL in Medical Registration: Review	MRI, CT, US, PET	TRE, DSC, MI, PSNR	Transformer-based unsupervised registration
14	Yuchong Gu et al.	2020	MedSRGAN	LUNA16 CT, Brain MRI	PSNR, SSIM, MOS	GAN–Transformer hybrids; federated/self-supervised SR
15	Zhihao Wang et al.	2021	DL for Image SR (Survey)	DIV2K, Set5, ImageNet	PSNR, SSIM, LPIPS	Efficient GAN/Transformer hybrids for SR
16	Yutong Xie et al.	2021	CoTr (CNN + Transformer Hybrid)	BCV Abdominal CT (30 scans)	Dice coefficient	Joint segmentation–SR framework
17	Xiangbin Liu et al.	2021	FCN, U-Net, GAN, DeepLab, V-Net	MRI, CT, X-ray, US	Dice, IoU, Accuracy	Transformer integration, self-supervision
18	Yue Yu et al.	2021	WFSAN (Wavelet Frequency Separation Attention Net)	Chest X-ray (Shenzhen, Montgomery)	PSNR, SSIM	Multi-level wavelet + Transformer design
19	Waqar Ahmad et al.	2022	Multi-Path Progressive GAN	DRIVE, STARE, BraTS, CAMUS	PSNR, SSIM	Lightweight diffusion-based GAN SR
20	Heng-Sheng Chao et al.	2022	SRCNN, VDSR, EDSR for Thin-Slice SR	Chest CT (35 studies)	PSNR, SSIM	Cross-modality SR; add Transformers

21	Bai Zhengyao & Tao Jinyu	2022	P3DSRNet (Pseudo-3D SR)	Kirby-21 MRI	PSNR, SSIM, FLOPs	Extend to multi-modal SR
22	Defu Qiu et al.	2023	Survey: Medical Image SR Algorithms	MRI, CT, PET, US	PSNR, SSIM, MSE	Hybrid DL-GAN frameworks; lightweight SR
23	Hujun Yang et al.	2023	Review: Deep Learning in Medical SR	MRI, CT, PET, US	PSNR, SSIM, LPIPS	Transformer-GAN integration; perceptual metrics
24	Walid El-Shafai et al.	2024	SISR Review for Medical Imaging	COVID-CT, DRIVE, BraTS	PSNR, SSIM, FSIM	Real-time SR; lightweight GAN-Transformer
25	Shizhuang Weng et al.	2025	DWTN + AMResNet	Hyperspectral Wheat Dataset	PSNR, SSIM, ACC	RefSR + segmentation-aware fusion

ML-Based Architectures

Early super-resolution (SR) in medical imaging relied on machine learning (ML) techniques such as dictionary learning, sparse coding, and regression-based mapping. These methods work well for structured anatomical patterns but often struggle with high variability and large-scale datasets.

1. Dictionary Learning:

Uses overcomplete dictionaries to reconstruct high-resolution (HR) patches from low-resolution (LR) inputs [33]. Widely used in early CT and MRI SR.

2. Sparse Representation:

Formulates SR as sparse coding in a high-resolution dictionary space, enabling detail enhancement even with limited data.

3. Regression-Based Methods:

Learn LR-to-HR mappings from paired patches (e.g., neighbor embedding, anchored regression).

Limitations:

- Not end-to-end
- High computational cost at inference
- Poor generalization across scanners or modalities

DL-Based Architectures

Deep learning (DL) revolutionized medical super-resolution (SR) by enabling end-to-end

learning of complex anatomical structures. Major architectural families include:

1. CNN-Based SR:

Examples include SRCNN, EDSR, and RCAN. These models are widely applied in CT, MRI, and ultrasound imaging.

2. GAN-Based SR:

SRGAN, ESRGAN, and pGAN improve perceptual quality and fine detail preservation [14].

3. Transformer-Based SR:

Models such as SwinIR and IPT introduce global context modeling for large-scale medical images [33].

4. GCN-Based SR:

Graph convolutional networks leverage structural relationships across anatomical regions or patches.

Advantages:

- End-to-end learning
- High PSNR/SSIM
- Better domain transfer and generalization

Hybrid ML-DL Approaches

Hybrid super-resolution (SR) models combine the interpretability of machine learning (ML) with the representational power of deep learning (DL). Examples include:

- Dictionary priors integrated with convolutional neural networks (CNNs) for stable reconstruction
- Sparse coding layers embedded within deep networks
- Model-based iterative reconstruction guided by deep priors

Lightweight and Real-Time Models

Clinical deployment of super-resolution (SR) models requires low latency and high computational efficiency. Lightweight architectures are designed to minimize parameters and memory requirements without compromising reconstruction quality.

Strategies:

- Depthwise separable convolutions (e.g., MobileSR)
- Knowledge distillation
- Efficient sub-pixel up sampling
- Model pruning and quantization

Network and Model Design Complexity

Modern super-resolution (SR) architectures such as Generative Adversarial Networks (GANs), Transformers, and Graph Neural Networks (GNNs) achieve state-of-the-art performance but introduce significant computational challenges. These include increased model size, complex training dynamics, and higher resource requirements.

Key Challenges:

- Large models are difficult to train with limited medical data
- Training instability in adversarial networks
- High memory and inference cost for 3D or volumetric SR

Lightweight super-resolution (SR) models and neural architecture search (NAS) are promising directions for reducing computational complexity while maintaining high reconstruction accuracy [75].

Domain Shift and Generalization

Super-resolution (SR) models trained on one dataset often underperform when applied to data from different scanners, institutions, or patient populations. This domain shift arises due to variations in imaging protocols, noise characteristics, and anatomical structures [73].

Examples:

- MRI field strength variation (1.5T vs 3T)
- Different CT reconstruction kernels
- Scanner manufacturer differences

Domain adaptation, self-supervised learning, and harmonization techniques are promising solutions to mitigate domain shift and improve generalization across diverse clinical environments [70].

Real-World Deployment Challenges

Even high-performing super-resolution (SR) models face significant challenges when transitioning from research to clinical deployment:

Regulatory compliance [83]: FDA and EMA approvals require extensive validation and reproducibility studies.

Integration: Seamless interfacing with PACS, RIS, and scanner systems is essential for workflow compatibility.

Real-time requirements: Clinical environments demand low-latency inference for time-critical decision-making.

Interpretability and trust: Clinicians require explainable SR outputs to ensure diagnostic reliability.

Bridging the gap between laboratory performance and real-world clinical adoption requires close collaboration among researchers, clinicians, and industry partners [79].

Future Research Directions

Although deep learning has significantly advanced medical image super-resolution (SR), the field remains far from clinical maturity. Future research should emphasize data-efficient learning, robustness, explainability, an

d practical deployment.

Self-Supervised and Unsupervised SR

Traditional supervised super-resolution (SR) methods rely on paired low-resolution (LR) and high-resolution (HR) datasets, which are often difficult to acquire in medical imaging. Self-supervised [24] and unsupervised SR approaches [94] eliminate the need for ground-truth HR images by learning from internal image statistics or domain-specific priors.

Potential Benefits:

- Works with unpaired or limited data
- Reduces annotation and acquisition cost
- Enables adaptation to rare modalities or institutions

Applications include MRI reconstruction without paired datasets and pathology image enhancement using internal learning.

Domain Adaptation

Medical super-resolution (SR) models often fail to generalize under domain shift. Domain adaptation and generalization approaches [96] aim to make SR models robust across different scanners, imaging protocols, and patient populations.

Strategies:

- Adversarial domain alignment
- Self-supervised feature harmonization
- Style transfer for cross-domain consistency

Domain-adapted SR models will facilitate reliable deployment in multi-center studies and low-resource clinical environments.

Federated and Privacy-Preserving SR

Patient privacy and data governance are major challenges for large-scale super-resolution (SR) model training. Federated learning (FL) enables collaborative model training across multiple institutions without requiring centralized data sharing [84][76].

Key Advantages:

- Preserves patient confidentiality

- Utilizes diverse data from multiple centers
- Improves generalization without sharing raw images

Future research should focus on integrating SR with federated learning and differential privacy techniques to support scalable and secure clinical deployment.

Multi-Task SR

Future super-resolution (SR) models are expected to integrate multiple related tasks such as segmentation, detection, or registration into a single framework. This multi-task learning approach can enhance both image reconstruction quality and downstream clinical performance [25][78].

Advantages:

- Shared features improve generalization
- Reduces inference time
- Enables end-to-end clinical workflows

Examples include joint SR with tumor segmentation in MRI or combining SR with registration for longitudinal imaging studies.

Efficient SR for Low-Resource Clinical Settings

Many hospitals, particularly in low-resource regions, lack access to high-end computing hardware. Designing efficient super-resolution (SR) architectures—such as lightweight CNNs, knowledge distillation, and quantized models—is crucial for practical, real-world adoption [73][81].

Future Focus:

- Edge and mobile deployment
- On-scanner real-time SR
- Energy-efficient inference

Explainable SR Models

To gain clinical trust, super-resolution (SR) models must be interpretable and transparent. Explainable SR (XSR) techniques enable radiologists to understand how image details are reconstructed and validated [73].

Key Directions:

- Uncertainty quantification and confidence maps
- Attribution and saliency visualization
- Integration of explainable AI (XAI) for regulatory approval

These approaches will accelerate the acceptance of SR algorithms in regulated clinical workflows.

Conclusion

Medical image super-resolution (SR) has evolved from conventional interpolation and sparse-representation techniques to advanced deep learning paradigms, significantly enhancing spatial fidelity, structural consistency, and diagnostic reliability across imaging modalities. By reconstructing high-resolution details from low-quality acquisitions, SR contributes to faster scans, reduced radiation exposure, and improved clinical interpretability in MRI, CT, ultrasound, digital pathology, and PET/SPECT.

This survey has provided a comprehensive synthesis of medical image SR methodologies, encompassing foundational concepts, network architectures, benchmark datasets, and evaluation metrics. It systematically reviewed the progression from early CNN-based frameworks to advanced GAN-, Transformer-, and graph-based architectures, highlighting their strengths, limitations, and clinical implications. Comparative analyses revealed the trade-offs between visual fidelity, computational cost, and clinical relevance, while emphasizing the growing importance of hybrid and lightweight designs for practical deployment.

Despite remarkable progress, several challenges persist—most notably, limited availability of annotated medical datasets, privacy restrictions, cross-domain variability, interpretability concerns, and the absence of standardized

evaluation protocols. Addressing these issues will be critical for bridging the gap between algorithmic innovation and real-world clinical integration.

Future research directions should emphasize:

- Self-supervised and unsupervised SR strategies to reduce dependence on paired datasets
- Domain adaptation and harmonization techniques for multi-center generalization
- Federated and privacy-preserving SR frameworks enabling secure data collaboration
- Multi-task learning to jointly optimize SR with segmentation, reconstruction, or diagnosis
- Lightweight and energy-efficient SR architectures for deployment in low-resource environments
- Explainable SR models that enhance clinician trust and regulatory transparency

Ultimately, the clinical translation of medical image SR will depend on cross-disciplinary collaboration among researchers, clinicians, and policymakers. Establishing unified benchmarking standards, rigorous validation pipelines, and interpretable SR frameworks will be essential for achieving regulatory acceptance and widespread adoption.

In summary, medical image SR represents a cornerstone of next-generation diagnostic imaging and precision medicine. With continuous advancements in algorithmic design, data governance, and clinical validation, SR is poised to become an indispensable component of future healthcare imaging ecosystems, enhancing both accessibility and quality of patient care.

References

1. G. Wang, H. Yu, and B. De Man, "An outlook on X-ray CT research and development," *Med. Phys.*, vol. 45, no. 11, pp. e1055–e1070, 2018.
2. C. H. McCollough, "Low-dose CT for the detection and classification of lung nodules," *Cancer Imaging*, vol. 15, no. 1, pp. 1–10, 2015.
3. D. J. Brenner and E. J. Hall, "Computed tomography—An increasing source of radiation exposure," *N. Engl. J. Med.*, vol. 357, no. 22, pp. 2277–2284, 2007.
4. M. Lustig, D. Donoho, and J. M. Pauly, "Sparse MRI: The application of compressed sensing for rapid MR imaging," *Magn. Reson. Med.*, vol. 58, no. 6, pp. 1182–1195, 2007.
5. T. L. Szabo, *Diagnostic Ultrasound Imaging: Inside Out*. Amsterdam, The Netherlands: Elsevier, 2004.
6. M. Mandal et al., "Deep learning for medical image super-resolution: A review," *Comput. Med. Imaging Graph.*, 2021.
7. Ebner et al., "Super-resolution reconstruction enables high-resolution MRI at reduced scan time," *Magn. Reson. Med.*, 2020.
8. S. Zhou et al., "Low-dose CT enhancement using deep learning-based super-resolution," *IEEE Trans. Med. Imaging*, 2021.
9. R. Keys, "Cubic convolution interpolation for digital image processing," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 29, no. 6, pp. 1153–1160, 1981.
10. J. Yang, J. Wright, T. Huang, and Y. Ma, "Image super-resolution via sparse representation," *IEEE Trans. Image Process.*, vol. 19, no. 11, pp. 2861–2873, 2010.
11. R. Zeyde, M. Elad, and M. Protter, "On single image scale-up using sparse-representations," in *Proc. Int. Conf. Curves Surf.*, 2010, pp. 711–730.
12. R. Timofte, V. De Smet, and L. Van Gool, "Anchored neighborhood regression for fast example-based super-resolution," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, 2013, pp. 1920–1927.
13. C. Dong, C. C. Loy, K. He, and X. Tang, "Image super-resolution using deep convolutional networks," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 38, no. 2, pp. 295–307, 2016.
14. C. Ledig, L. Theis, F. Huszár, J. Caballero, A. Cunningham, A. Acosta, A. Aitken, A. Tejani, J. Totz, Z. Wang, and W. Shi, "Photo-realistic single image super-resolution using a generative adversarial network," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, 2017, pp. 4681–4690.
15. B. Lim, S. Son, H. Kim, S. Nah, and K. M. Lee, "Enhanced deep residual networks for single image super-resolution," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. Workshops (CVPRW)*, 2017, pp. 136–144.
16. J. Chen et al., "Trans-SR: Transformer-based medical image super-resolution," *Med. Image Anal.*, 2022.
17. Y. He et al., "Graph neural network enhanced super-resolution for medical imaging," *IEEE Trans. Med. Imaging*, 2023.
18. Y. Chen et al., "Low-dose CT with a residual encoder-decoder convolutional neural network," *IEEE Trans. Med. Imaging*, 2018.

19. S. C. Park, M. K. Park, and M. G. Kang, "Super-resolution image reconstruction: A technical overview," *IEEE Signal Process. Mag.*, vol. 20, no. 3, pp. 21–36, 2003.
20. R. Y. Tsai and T. S. Huang, "Multiframe image restoration and registration," *Adv. Comput. Vis. Image Process.*, vol. 1, pp. 317–339, 1984.
21. D. Glasner, S. Bagon, and M. Irani, "Super-resolution from a single image," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, 2009.
22. S. Farsiu, D. Robinson, M. Eand high-frequency ultrasound volumeslad, and P. Milanfar, "Fast and robust multiframe super resolution," *IEEE Trans. Image Process.*, vol. 13, no. 10, pp. 1327–1344, 2004.
23. D. Capel, *Image Mosaicing and Super-Resolution*. Berlin, Germany: Springer, 2004.
24. A. Shocher, N. Cohen, and M. Irani, "'Zero-shot' super-resolution using deep internal learning," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, 2018.
25. K. Zhang, W. Zuo, and L. Zhang, "Deep plug-and-play super-resolution for arbitrary blur kernels," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, 2019.
26. K. Zhang et al., "Image restoration using very deep convolutional encoder-decoder networks with symmetric skip connections," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, 2018.
27. Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From T. Tong, G. Li, X. Liu, and Q. Gao, "Image super-resolution using dense skip connections," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, 2017.error visibility to structural similarity," *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 600–612, 2004.
28. J. Gu et al., "Losses and metrics for image restoration: From fidelity to perceptual quality," *arXiv preprint arXiv:2104.14085*, 2021.
29. R. Zhang, P. Isola, A. A. Efros, E. Shechtman, and O. Wang, "The unreasonable effectiveness of deep features as a perceptual metric," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, 2018, pp. 586–595.
30. J. Kim, J. K. Lee, and K. M. Lee, "Accurate image super-resolution using very deep convolutional networks," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, 2016, pp. 1646–1654.
31. W. Shi, J. Caballero, F. Huszár, J. Totz, A. Aitken, R. Bishop, D. Rueckert, and Z. Wang, "Real-time single image and video super-resolution using an efficient sub-pixel convolutional neural network," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, 2016, pp. 1874–1883.
32. W.-S. Lai, J.-B. Huang, N. Ahuja, and M.-H. Yang, "Deep Laplacian pyramid networks for fast and accurate super-resolution," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, 2017.
33. J. Liang et al., "SwinIR: Image restoration using Swin Transformer," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, 2021.
34. Y. Rivenson, Y. Wu, H. Wang, Z. Wang, and A. Ozcan, "Deep learning microscopy," *Optica*, vol. 5, no. 4, pp. 352–360, 2018.
35. C. E. Duchon, "Lanczos filtering in one and two dimensions," *J. Appl. Meteorol.*, vol. 18, no. 8, pp. 1016–1022, 1979.

-
36. J. A. Parker, R. V. Kenyon, and D. E. Troxel, "Comparison of interpolating methods for image resampling," *IEEE Trans. Med. Imaging*, vol. 2, no. 1, pp. 31–39, 1983.
 37. T. M. Lehmann, C. Gönnér, and K. Spitzer, "Survey: Interpolation methods in medical image processing," *IEEE Trans. Med. Imaging*, vol. 18, no. 11, pp. 1049–1075, 1999.
 38. P. Thévenaz, T. Blu, and M. Unser, "Interpolation revisited," *IEEE Trans. Med. Imaging*, vol. 19, no. 7, pp. 739–758, 2000.
 39. M. Unser, "Splines: A perfect fit for signal and image processing," *IEEE Signal Process. Mag.*, vol. 16, no. 6, pp. 22–38, 1999.
 40. M. Aharon, M. Elad, and A. Bruckstein, "K-SVD: An algorithm for designing overcomplete dictionaries for sparse representation," *IEEE Trans. Signal Process.*, vol. 54, no. 11, pp. 4311–4322, 2006.
 41. Y. Chen, Y. Li, L. Zhang, and Q. Hu, "Sparse representation and dictionary learning for medical image super-resolution," *IEEE Trans. Med. Imaging*, vol. 32, no. 12, pp. 2179–2189, 2013.
 42. W. T. Freeman, T. R. Jones, and E. C. Pasztor, "Example-based super-resolution," *IEEE Comput. Graph. Appl.*, vol. 22, no. 2, pp. 56–65, 2002.
 43. H. Chang, D.-Y. Yeung, and Y. Xiong, "Super-resolution through neighbor embedding," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, 2004, pp. 275–282.
 44. R. Timofte, V. De Smet, and L. Van Gool, "A+: Adjusted anchored neighborhood regression for fast super-resolution," in *Proc. Asian Conf. Comput. Vis. (ACCV)*, 2014, pp. 111–126.
 45. S. Schulter, C. Leistner, and H. Bischof, "Fast and accurate image upscaling with super-resolution forests," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, 2015, pp. 3791–3799.
 46. Q. Yang et al., "Low-dose CT image enhancement using patch-based sparse representation," *Phys. Med. Biol.*, vol. 58, no. 16, p. 5803, 2013.
 47. S. T. Roweis and L. K. Saul, "Nonlinear dimensionality reduction by locally linear embedding," *Science*, vol. 290, no. 5500, pp. 2323–2326, 2000.
 48. M. Belkin and P. Niyogi, "Laplacian eigenmaps for dimensionality reduction and data representation," *Neural Comput.*, vol. 15, no. 6, pp. 1373–1396, 2003.
 49. J. Ma et al., "Manifold alignment for super-resolution image reconstruction," *Inf. Fusion*, vol. 23, pp. 124–131, 2015.
 50. J.V. Manjón, P. Coupé, A. Buades, D. L. Collins, and M. Robles, "New methods for MRI denoising based on sparseness and self-similarity," *Med. Image Anal.*, vol. 12, no. 5, pp. 497–508, 2010.
 51. A. Zog, A. Carass, and J. L. Prince, "Patch-based super-resolution of brain MRI using coupled sparse dictionaries," *IEEE Trans. Med. Imaging*, vol. 33, no. 10, pp. 2003–2016, 2014.
 52. M. Elad and M. Aharon, "Image denoising via sparse and redundant representations over learned dictionaries," *IEEE Trans. Image Process.*, vol. 15, no. 12, pp. 3736–3745, 2006.
 53. M. Protter, M. Elad, H. Takeda, and P. Milanfar, "Generalizing the nonlocal-means to super-resolution reconstruction," *IEEE Trans. Image Process.*, vol. 18, no. 1, pp. 36–51, 2009.
 54. J. Zhou et al., "Patch-based super-resolution reconstruction of CT

-
- images,” *Phys. Med. Biol.*, vol. 56, pp. 2475–2490, 2011.
55. Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, “Gradient-based learning applied to document recognition,” *Proc. IEEE*, vol. 86, no. 11, pp. 2278–2324, 1998.
 56. C. Dong, C. C. Loy, K. He, and X. Tang, “Learning a deep convolutional network for image super-resolution,” in *Proc. Eur. Conf. Comput. Vis. (ECCV)*, 2014, pp. 184–199.
 57. V. Nair and G. E. Hinton, “Rectified linear units improve restricted Boltzmann machines,” in *Proc. Int. Conf. Mach. Learn. (ICML)*, 2010.
 58. K. He, X. Zhang, S. Ren, and J. Sun, “Delving deep into rectifiers: Surpassing human-level performance on ImageNet classification,” in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, 2015.
 59. S. Ioffe and C. Szegedy, “Batch normalization: Accelerating deep network training by reducing internal covariate shift,” in *Proc. Int. Conf. Mach. Learn. (ICML)*, 2015.
 60. . He, X. Zhang, S. Ren, and J. Sun, “Deep residual learning for image recognition,” in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, 2016.
 61. G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, “Densely connected convolutional networks,” in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, 2017.
 62. Y. Zhang, K. Li, K. Li, L. Wang, B. Zhong, and Y. Fu, “Image super-resolution using very deep residual channel attention networks,” in *Proc. Eur. Conf. Comput. Vis. (ECCV)*, 2018.
 63. J. Hu, L. Shen, and G. Sun, “Squeeze-and-excitation networks,” in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, 2018.
 64. I. Goodfellow et al., “Generative adversarial nets,” in *Proc. Adv. Neural Inf. Process. Syst. (NeurIPS)*, 2014.
 65. M. Mardani et al., “Deep generative adversarial neural networks for compressive sensing MRI,” *IEEE Trans. Med. Imaging*, 2018.
 66. J.-Y. Zhu, T. Park, P. Isola, and A. A. Efros, “Unpaired image-to-image translation using CycleGAN,” in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, 2017.
 67. A. Vaswani et al., “Attention is all you need,” in *Proc. Adv. Neural Inf. Process. Syst. (NeurIPS)*, 2017.
 68. H. Chen et al., “Pre-trained image processing transformer,” in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, 2021.
 69. I. Feigin et al., “Deep learning for ultrasound super-resolution,” *Ultrasound Med. Biol.*, 2018.
 70. B. Glocker et al., “Machine learning with multi-site imaging data: An empirical study on the impact of scanner effects,” *Med. Image Anal.*, 2019.
 71. A. Greenbaum et al., “Wide-field computational imaging of pathology slides using lens-free on-chip microscopy,” *Sci. Transl. Med.*, 2014.
 72. W. Han, S. Chang, D. Liu, and M. Yu, “Image super-resolution via dual-state recurrent networks,” in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, 2018.
 73. A. Howard et al., “Searching for MobileNetV3,” in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, 2019.
 74. X. Hu et al., “Meta-SR: A magnification-arbitrary network for super-resolution,” in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, 2019.

-
75. S. Huang et al., "Self-supervised learning for domain generalization in medical imaging," *Med. Image Anal.*, 2022.
 76. P. Isola, J.-Y. Zhu, T. Zhou, and A. A. Efros, "Image-to-image translation with conditional adversarial networks," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, 2017.
 77. J. Johnson, A. Alahi, and L. Fei-Fei, "Perceptual losses for real-time style transfer and super-resolution," in *Proc. Eur. Conf. Comput. Vis. (ECCV)*, 2016.
 78. D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," in *Proc. Int. Conf. Learn. Represent. (ICLR)*, 2015.
 79. G. A. Kaissis et al., "End-to-end privacy preserving deep learning on multi-institutional medical imaging," *Nat. Mach. Intell.*, 2021.
 80. T. N. Kipf and M. Welling, "Semi-supervised classification with graph convolutional networks," in *Proc. Int. Conf. Learn. Represent. (ICLR)*, 2017.
 81. S. Leclerc et al., "Deep learning for segmentation using the CAMUS dataset," *Med. Image Anal.*, 2019.
 82. H. Lee et al., "Regulatory perspectives on AI/ML-based medical devices," *NPJ Digit. Med.*, 2022.
 83. B. H. Menze et al., "The multimodal brain tumor image segmentation benchmark (BRATS)," *IEEE Trans. Med. Imaging*, 2015.
 84. S. J. Sheller et al., "Federated learning in medicine: Facilitating multi-institutional collaborations without sharing patient data," *Sci. Rep.*, 2020.
 85. R. Timofte et al., "NTIRE 2016 challenge on single image super-resolution: Methods and results," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. Workshops (CVPRW)*, 2016.
 86. J. Loshchilov and F. Hutter, "SGDR: Stochastic gradient descent with warm restarts," in *Proc. Int. Conf. Learn. Represent. (ICLR)*, 2016.
 87. T. Elsken, J. H. Metzen, and F. Hutter, "Neural architecture search: A survey," *J. Mach. Learn. Res. (JMLR)*, 2019.
 88. S. H. Park et al., "CT image super-resolution using GAN for low-dose enhancement," *Med. Phys.*, 2020.
 89. A. J. Reader and J. Verhaeghe, "4D image reconstruction for emission tomography," *Phys. Med. Biol.*, vol. 59, no. 22, pp. R371–R418, 2014.
 90. J. N. Weinstein et al., "The Cancer Genome Atlas Pan-Cancer analysis project," *Nat. Genet.*, 2013.
 91. K. Clark et al., "The Cancer Imaging Archive (TCIA): Maintaining and operating a public information repository," *J. Digit. Imaging*, 2013.
 92. Y. Zhang et al., "A multi-task learning framework for image restoration," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, 2019.
 93. K. Zhang et al., "Plug-and-play image restoration with deep denoiser prior," *IEEE Trans. Pattern Anal. Mach. Intell.*, 2020.
 94. X. Wang et al., "Unsupervised real-world image super-resolution," *Int. J. Comput. Vis. (IJCV)*, 2021.
 95. G. Wang et al., "Deep learning for tomographic image reconstruction," *Nat. Mach. Intell.*, 2020.
 96. X. Zhou et al., "Domain adaptation in medical image analysis: A survey," *Med. Image Anal.*, 2022.
 97. J. Zhou et al., "Hybrid dictionary and CNN based SR for medical imaging," *Med. Image Anal.*, 2021.
 98. Y. Zhou et al., "CT image super-resolution using GAN for low-dose enhancement," *Med. Phys.*, 2020.

-
99. N. Arun et al., “Assessing the trustworthiness of saliency maps for localizing abnormalities in medical imaging,” *Nat. Mach. Intell.*, 2021.
 100. Y. Blau and T. Michaeli, “The perception-distortion tradeoff,” in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, 2018.
 101. T. Tong, G. Li, X. Liu, and Q. Gao, “Image super-resolution using dense skip connections,” in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, 2017.