

Priors in Medical Image Analysis: A Review

Abdulloh Umarov*

Department of Biomedical Engineering, Tashkent State Technical University, Tashkent, Uzbekistan

*Corresponding author: Abdulloh Umarov, Department of Biomedical Engineering, Tashkent State Technical University, Tashkent, Uzbekistan,
E-mail: scorpionking8927@gmail.com

Received date: Aug 26, 2026; Accepted date: Sep 01, 2026; Published date: Sep 26, 2026

Citation: Umarov A (2026) Artificial Intelligence–Driven Intelligent Project Management: An Integrated Framework for Planning, Scheduling and Control in Engineering and Construction Projects. J Eng Artif Intell Vol.2 No.2: 53.

Abstract

Medical imaging tasks are frequently limited by data that is either insufficient or excessive. In tomographic reconstruction, interventional fluoroscopy and endoscopic imaging, the available measurements are incomplete: Projections are few, the angular range is narrow, only one or two views are acquired or the observations are corrupted by physical effects that the imaging model does not account for. In digital pathology the situation is reversed, since a single whole-slide image may contain 109-1010 pixels and 104-105 usable tiles and therefore cannot be processed at native resolution. This review examines both situations and notes that each is addressed by the same fundamental mechanism, namely the introduction of a prior that constrains the space of admissible solutions. Over the past three decades the necessity of the prior has not changed, but its form has: Analytic smoothness penalties were succeeded by hand-designed variational regularizers and patch-based self-similarity models and subsequently by learned regularizers embedded in unrolled iterations, untrained network architectures, generative score models, attention distributions over gigapixel fields and large-scale foundation encoders whose embeddings provide task-agnostic representations. We review this development across five application areas image restoration, sparse-view and limited-angle tomography, biplane interventional imaging, endoscopic vision and computational pathology and summarise the limitations that remain unresolved, including hallucination of structure in the undetermined null space, poorly calibrated uncertainty, covariate shift across scanners and institutions and evaluation metrics that correlate weakly with diagnostic performance.

Keyword: Inverse problems; Total variation; Sparse-view CT; Limited-angle tomography; Deep image prior; Biplane angiography; Multiple-instance learning; Whole-slide imaging; Computational pathology; Domain adaptation

Introduction

Every medical imaging pipeline is an attempt to recover a clinically meaningful quantity from measurements that do not determine it. This is uncontroversial for tomography, where the statement is literal: Reconstructing a volume from a finite number of line integrals is an underdetermined linear inverse problem and reducing the number of projections to limit radiation dose makes the deficiency worse in a precisely characterizable way. It is equally true, if less often stated in those terms, for interventional imaging, where a three-dimensional vascular tree or a device pose must be inferred from two projections; for endoscopy, where a near-coaxial point source on a wet, deformable, non-Lambertian mucosa produces radiance that systematically violates the photometric assumptions of every downstream algorithm; and, in a way that initially looks like the opposite problem, for digital pathology,

where the acquired data so vastly exceeds what any model can consume that the practical task becomes deciding what to discard. A common observation across all four settings is that the deficiency is addressed by the same operation. A prior is imposed a statement about what admissible images look like, which regions carry diagnostic signal or how the observation was corrupted and this prior selects one solution from the family that the data alone permits. Tikhonov's regularization of ill-posed problems [1] is the earliest general statement of the strategy; the Rudin–Osher–Fatemi total variation functional [2] is its most consequential single instance.

Nothing in the subsequent thirty years has removed the need for the prior. What has changed, repeatedly and in one consistent direction, is where the prior lives: It has migrated out of the mathematician's choice of functional and into data.

This migration can be described in five stages, which are set out explicitly here because they recur independently in each of the domains surveyed. In the first stage the prior is an analytic penalty on the image, chosen for the geometry it favours: quadratic smoothness, bounded variation, piecewise smoothness with an explicit edge set [3], anisotropic diffusion along level lines [4]. In the second the prior becomes a statement about patches rather than derivatives: Images are self-similar, so a region is predicted by other regions of the same image [5,6] or sparse in a dictionary learned from data rather than fixed in advance [7]. In the third the prior is learned discriminatively, as a direct map from corrupted to clean [8] or as a regularizer unrolled inside a physics-consistent iteration [9,10]. In the fourth the prior is decoupled from the forward operator once more as an untrained network architecture [11] or a score-based generative model [12] restoring the operator-agnostic flexibility that supervised networks had surrendered. In the fifth, currently unfolding in computational pathology, the prior is a foundation-model embedding: A frozen encoder trained on billions of tiles supplies a representation in which downstream clinical tasks become nearly linear [13,14]. The author's own work spans this period and is used here to illustrate the progression rather than to assert priority.

It begins in the classical variational era with total-variation denoising and hybrid median filtering for X-ray restoration [15,16], passes through algebraic reconstruction and few-view CT [17,18], through projection-domain interpolation as a means of manufacturing the views that were never acquired [19,20], into interventional biplane imaging for vascular reconstruction and magnetic microrobot pose estimation [21,22] and arrives in the learned-prior era with deep-prior limited-angle tomography, prior-based sparse-view CBCT, endoscopic specular-region handling and gigapixel whole-slide compression and weakly supervised diagnosis [23-29]. The trajectory is not unusual; it is what the field did, seen from one vantage point. The review is organized as follows. Section 2 establishes the foundational problems: missing-region completion, multi-view geometry and corrupted-channel inference whose formal machinery the later medical work inherits. Section 3 treats hand-designed priors for restoration. Section 4 treats sparse-view and limited-angle tomography, in which the contribution of the prior is most apparent. Section 5 treats interventional biplane imaging, where the prior must additionally support real-time control. Section 6 treats endoscopic vision as a corrupted-observation problem requiring explicit photometric models. Section 7 treats the overwhelming regime of computational pathology. Section 8 addresses the deployment and evaluation constraints that

determine whether any of this reaches a patient and Section 9 states what remains unsolved.

Foundational Concepts: Missing Data, Multi-View Geometry and Noisy Channels

Three lines of work outside medical imaging supply the formal apparatus that the rest of this review uses and it is worth being explicit about the correspondences rather than gesturing at them.

Missing-region completion

Image inpainting is the canonical missing-data problem: Recover a function on a masked domain from its values outside the mask, plus a prior. Bertalmio et al. [30] solved it by propagating isophotes into the unobserved region under a third-order transport equation variational completion in the strict sense and formally the same operation as recovering an unsampled angular wedge in Radon space, which is likewise a mask. Efros and Leung [31] attacked it instead by conditional non-parametric sampling, predicting a pixel from the statistics of neighbourhoods elsewhere in the image; Efros and Freeman [32] made this practical through patch quilting and Criminisi et al. [33] added a confidence-ordered traversal so that the most constrained parts of the mask are filled first a priority schedule that reappears, in different clothing, as data-consistency weighting inside iterative reconstruction. Our early work on texture-synthesis-based region duplication [34] sits squarely in this exemplar-based tradition. Pathak et al. [35] then replaced the hand-designed completion prior with a learned generative one, which is the transition this review traces, arriving in the general vision literature roughly a decade before it dominated tomography.

Multiple views

Longuet-Higgins [36] established that two projections of a rigid scene are related by a bilinear epipolar constraint and that scene structure and relative geometry can be recovered jointly from correspondences. This is the entire mathematical basis of biplane angiographic reconstruction and Metz and Fencil [37] stated the medical version directly: Three-dimensional structure is recoverable from biplane radiographs without prior knowledge of the inter-view relationship. Tomasi and Kanade [38] contributed the complementary insight that a measurement matrix of tracked points under orthography has rank three, so structure and motion follow from a low-rank factorization a structural prior that resurfaces as low-rank-plus-sparse modelling in dynamic MRI and motion-compensated cone-beam CT. Schönberger and Frahm

[39] turned incremental bundle adjustment into robust practice. Our early exploration of three-dimensional structure recovery from forward camera motion [40] and of motion-vector estimation in transparent, low-texture sequences [41] belongs to this lineage and the transparency case is instructive: Superimposed layers break the brightness-constancy assumption in exactly the way that projective X-ray imaging does, since every fluoroscopic frame is a semi-transparent composite rather than a surface observation.

Noisy channels

Shannon [42] fixed the ceiling: What is recoverable from a corrupted observation is bounded by the channel's capacity, which in imaging terms bounds what any reconstruction method can extract at a given dose or exposure. This is the correct frame for evaluating dose-reduction claims and it is why noise-model-specific likelihoods Poisson at low photon counts, Rician in magnitude MRI are not implementation details but the data-fidelity half of every MAP formulation in this review. Communication engineering also supplies the transform-domain habit of mind: The DFT-based multicarrier scheme of Weinstein and Ebert [43], elaborated in Bingham [44] treatment, equalizes frequency-selective intersymbol interference per subcarrier, which is the discrete-Fourier analogue of modulation-transfer-function deconvolution in projection radiography. Our own analysis of intersymbol interference reduction over an additive-white-Gaussian-noise channel [45] is an instance of this reasoning and the habit it instils model the corruption explicitly, then invert it under a prior is the one that transfers. A fourth, more prosaic foundation deserves mention because it constrains deployment rather than theory. Distributed clinical sensing inherits the energy, bandwidth and topology limits catalogued for wireless sensor networks [46], together with their threat model: Nodes can be compromised and inference pipelines that assume integrity are unsafe. Hierarchical intrusion detection for such networks [47] and the broader taxonomy of sensor-network intrusion detection [48], define the integrity assumptions that any bedside or wearable medical inference fabric silently relies on. We return to this in Section 8.

Hand-Designed Priors: Restoration and Denoising

The first medical application of the prior is restoration and the literature here is the cleanest available illustration of the general argument, because the priors are explicit and their failure modes are visible. Tikhonov [1] stabilized ill-posed inversion with a quadratic penalty on the solution's derivatives, at the cost of blurring precisely the discontinuities

that carry diagnostic meaning. Rudin et al. [2] replaced the squared gradient with its L1 norm, admitting piecewise-constant solutions and preserving edges and in doing so created the single most reused prior in medical imaging. Perona and Malik [4] encoded the same edge preference as a diffusion process whose conductance falls at high gradient magnitude and Mumford and Shah [3] as a free-discontinuity functional with an explicit edge set. All three are nonsmooth or nonconvex, which made the subsequent progress algorithmic rather than conceptual: Half-quadratic minimization introduces auxiliary variables that decouple the data term from the regularizer and reduce each step to a linear solve plus a pointwise nonlinearity [49], split Bregman iteration [50] and the Chambolle-Pock primal-dual algorithm [51] handle composite TV objectives at scale and accelerated proximal gradient methods [52] supply the convergence rates.

Our contributions in this period were to the practical behaviour of these solvers on radiographic data. Total-variation denoising applied directly to X-ray images [15] established the baseline; an L1-fidelity variant addressed the heavy-tailed, non-Gaussian corruption that X-ray detectors actually produce [53]; and a variable half-quadratic penalty parameter [54] attacked the practical bottleneck of half-quadratic schemes, whose convergence rate depends sharply on a penalty weight that is conventionally fixed. Continuation on that parameter loose early, tight late recovers most of the accuracy of a well-tuned static choice at a fraction of the iteration count, which matters when restoration is a preprocessing stage inside a reconstruction loop rather than an endpoint. Total variation alone, however, mishandles impulsive corruption, because a sparse set of extreme outliers is cheaper for the TV functional to interpolate than to reject. The classical remedy is rank-order filtering median filters and their adaptive-window extensions [55] and the correct combination is two-phase: Detect the impulse-corrupted pixels with a median-type detector, then restore only those under an edge-preserving variational penalty [56]. Our hybrid median filter combined with total-variation minimization [16] is an instance of this hybrid design applied to medical radiography, where mixed Gaussian-impulse noise is the realistic model rather than a textbook curiosity and were treating the two components with a single functional demonstrably degrades both.

The prior's support then widened from local derivatives to patches. Non-local means [5] predicts a pixel from perceptually similar neighbourhoods anywhere in the image; BM3D [6] groups similar patches into a three-dimensional stack and exploits sparsity in a transform of that stack and remained the strongest non-learned denoiser for a decade; K-SVD [7]

learned the sparsifying dictionary from the data itself, which is the first genuine step from designed to learned priors. Discriminative learning then dispensed with the explicit prior: DnCNN [8] learns the residual map directly and handles blind noise levels and windowed-attention architectures such as SwinIR [57] replaced convolutional locality with learned long-range interactions. The synthesis that closed this line is plug-and-play regularization [58]. Because variable splitting isolates a proximal step that is formally just a denoiser, any denoiser BM3D, DnCNN, a diffusion model can be inserted as the prior inside a model-based reconstruction, with the forward operator retained exactly. This is the mechanism by which the restoration literature of Section 3 becomes the reconstruction literature of Section 4 and the boundary between restoration and reconstruction has correspondingly become less distinct.

Sparse-View and Limited-Angle Tomography: The Role of Regularization

Tomographic reconstruction is the setting in which the prior's contribution is largest and most easily misread as information. Analytic inversion by filtered backprojection assumes dense angular sampling; when projections are decimated to reduce dose or the gantry cannot complete an arc, the resulting linear system is underdetermined and the least-squares solution is non-unique. Algebraic methods were introduced precisely to admit such geometries: The row-action scheme of Gordon et al. [59] reformulated reconstruction as an iterative linear solve, Gilbert [60] traded convergence speed for noise robustness by updating simultaneously, Andersen and Kak [61] introduced the ray-weighted simultaneous update that remains the standard inner solver, Wang et al. [62] supplied ordered-subset acceleration and Shepp and Vardi [63] recast the problem statistically with an explicit Poisson likelihood. None of these resolves non-uniqueness; they only make the solution set computationally accessible.

The prior enters as a constraint on that set. Sidky et al. [64] and Sidky and Pan [65] established the reference formulation: Alternate projection onto the data-consistency set with descent on total variation, an algorithm adaptive-steepest-descent projection onto convex sets, ASD-POCS that made few-view and limited-angle cone-beam reconstruction practically viable and remains the baseline two decades later. Chen et al. [66] generalized the idea in an important direction with prior-image constrained compressed sensing, replacing the generic sparsity norm with a patient-specific prior image, which is conceptually the bridge to everything learned that followed: The prior stopped being a statement about images in general

and became a statement about this anatomy. The limits of this strategy are known exactly, although this point is not always made explicit in the deep-learning literature. Friel and Quinto [67] showed by microlocal analysis that limited-angle data determines only those wavefront directions lying within the acquired angular range. Edges oriented outside that range are not poorly estimated; they are not determined at all and any method that renders them does so from the prior alone. The characteristic streak artifacts of limited-angle reconstruction are likewise predictable from the geometry rather than incidental. Any claim that a reconstruction method "recovers" such structure is a claim about the prior's accuracy, not about the measurement.

Our work in this area followed the standard path from solver engineering to learned priors. A modified simultaneous iterative reconstruction technique [18] addressed the convergence-quality trade-off in SIRT-family solvers under reduced projection counts; a modified conjugate-gradient total-variation algorithm [68] and an explicit total-variation CT formulation applied the variational machinery of Section 3 inside the reconstruction loop; and a few-view formulation [17] characterized the degradation as projections are removed. A distinct strand pursued a different tactic: Rather than strengthening the prior on the image, synthesize the missing projections. Interpolating in the projection domain to manufacture intermediate views before reconstruction [19,20] trades an ill-posed reconstruction for a better-posed one plus an interpolation error, which is favourable when the angular gap is small and the object is smooth in the sinogram's angular direction. The learned version of exactly this idea deep sinogram synthesis for sparse-view CT [69] became a substantial subfield and the comparison between the two is a clean illustration of what learning contributes: Not a new problem formulation, but a far better interpolant. Deep learning then entered reconstruction along three axes that are still distinguishable. Image-domain post-processing learns a map from artifact-laden FBP output to a clean image, exemplified by FBPCNN [70] and, for the low-dose rather than sparse-view case, RED-CNN [71] and its adversarial successor WGAN-VGG [72], the latter explicitly trading pixelwise fidelity for a noise texture radiologists prefer. Unrolled optimization embeds a learned regularizer inside a physics-consistent iteration, as in ADMM-Net [9] and LEARN [10], retaining the forward operator and hence much of the robustness that pure post-processing loses. Full domain-transform learning, as in AUTOMAP [73], discards the analytic operator entirely, which maximizes flexibility and minimizes guarantees. The most recent shift returns to operator-agnostic priors. The deep image prior demonstrated

that an untrained convolutional architecture, fitted to a single measurement, regularizes by architecture alone, requiring no training corpus a property Baguer et al. [11] exploited for CT.

Our deep-prior formulation of limited-angle tomography [25] applies this in the geometry where the null space is largest and the absence of a training distribution is most valuable, since supervised networks trained on one angular range generalize poorly to another. In parallel, we found that the classical framework still has room: Augmenting ASD-POCS with an explicit structural prior for artifact suppression in sparse-view cone-beam CT [23] improves structural preservation without abandoning the convergence behaviour and data consistency that make the classical algorithm auditable a consideration that matters more in a regulated setting than benchmark tables suggest. Score-based generative reconstruction [12] is the strongest current general-purpose prior, decoupling a single learned image distribution from the sampling geometry so that one model serves many acquisition protocols, at the cost of slow sampling and a well-documented tendency to produce plausible structure in the null space. Evaluation across all of these rests on a small number of shared resources, principally the AAPM Low Dose CT Grand Challenge and its task-based assessment protocol [74] and on structural rather than error-based fidelity metrics [62] whose relationship to diagnostic performance remains loose.

Interventional Imaging: Few-View Reconstruction and Device Tracking

Interventional imaging compresses the tomographic problem to its extreme. Two projections sometimes one must yield a three-dimensional vascular geometry or a device pose, at frame rates compatible with control, under a dose budget constrained by documented deterministic skin injury in prolonged fluoroscopically guided procedures [75]. Digital subtraction angiography [76] supplies the original prior in this domain: A mask image acquired before contrast arrival is subtracted from the filled sequence, making vessels observable in projections where they would otherwise be lost in overlying anatomy. Subtraction is a prior in disguise it assumes the anatomy is static between mask and fill and it fails exactly when that assumption does. Three-dimensional recovery from biplane views rests on the geometric results of Section 2, made medical by Metz and Fencil [37] and turned into a working pipeline by Chen and Carroll [77], whose key move was to add an anatomical prior: Vessels are thin, connected, tree-structured curves, so dense stereo correspondence collapses to epipolar matching of centrelines.

Blondel et al. [78] extended this by treating cardiac motion itself as the prior that substitutes for simultaneous views in a rotational sequence and Çimen et al. [79] catalogued the resulting design space, in which nearly every method combines geometry estimation, correspondence and a curve or intensity regularizer. The feature extractors carry priors of their own: Hessian-eigenvalue vesselness [80] encodes tubularity analytically, while learned encoder-decoder segmentation [81], benchmarked in the tradition established by DRIVE [82], replaced the filter bank with an appearance model robust to subtraction artifacts, vessel overlap and low-dose noise. Registration of a pre-interventional volume to intra-procedural projections [83] is the same logic applied at the level of whole anatomy and real-time device segmentation in fluoroscopy [84] supplies the device-side counterpart. Our contribution here was to automate the biplane pipeline end to end by combining learned segmentation with the geometric constraint rather than treating them as separate stages: Three-dimensional vessel reconstruction from bi-plane angiographic projections driven by deep segmentation [22] uses the network to establish reliable centreline correspondence under exactly the conditions foreshortening, overlap, low contrast where Hessian-based extraction becomes unreliable, then closes the geometry analytically. Recent neural implicit approaches [85] push further by encoding the shape prior in the representation itself, dispensing with explicit correspondence and their trade-off is instructive: Greater robustness, less auditability.

The most demanding application of few-view pose estimation is the control of magnetically actuated devices. Untethered microrobots for minimally invasive intervention [86] are steered by external electromagnetic systems whose control laws require metric pose in five or six degrees of freedom at high rate the OctoMag platform [87] is the canonical instance and the imaging requirement it imposes is unusually explicit, since the controller consumes pose directly rather than an image. Comparative surveys of microrobot localization across X-ray, ultrasound, MRI and photoacoustic modalities [88] make the trade-off plain: Only X-ray offers the temporal resolution and penetration depth that vascular navigation needs and only at a dose cost, which is precisely the argument for biplane rather than tomographic acquisition. Our real-time microrobot posture recognition system using a biplane X-ray platform for external electromagnetic actuation [21] addressed this directly, recovering orientation as well as position from two projections at control rates; recent *ex vivo* validation of magnetically actuated intravascular robots under clinical angiography [89] indicates the direction of travel toward the interventional suite.

Endoscopic Vision: Photometric Corruption and Its Correction

Endoscopy inverts the emphasis. The data is not sparse video is abundant but the observation model that downstream algorithms assume is false in a structured, exploitable way. A near-coaxial point source illuminating wet mucosa produces radiance that is a mixture of diffuse and specular components and Shafer [90] dichromatic reflection model states the mixture explicitly: The specular term carries the illuminant's chromaticity rather than the surface's and therefore carries no geometric information about the tissue. Separating the components from a single image is an ill-posed per-pixel unmixing closed by that physical prior, solved classically by iterating on a specular-free representation [91]. Once detected, specular pixels are best understood as missing data, which returns us to Section 2: Endoscopic highlight handling is segmentation followed by inpainting [92] and the modern learned version exploits temporal redundancy across frames and importantly is evaluated on downstream correspondence quality rather than on inpainting fidelity [93]. This matters because the corruption propagates upward into geometry. Handcrafted local features assume brightness constancy: SIFT [94], SURF [95] and ORB [96] the latter the front end of the monocular SLAM systems [97] on which much endoscopic tracking is built all detect structures that, on specular mucosa, are artifacts of the light source's position rather than covariant properties of the scene and the affine-covariance evaluation protocol of Mikolajczyk et al. [98] quantifies how badly repeatability degrades on low-texture, illumination-varying surfaces. Our work on specular region detection coupled with covariant feature extraction [26] treats these as one problem rather than two: Identifying specular regions is worthwhile chiefly because it tells the feature stage which detections to distrust and the correct output is a feature set with reliability annotations rather than a cosmetically cleaner image.

Learned detectors such as SuperPoint [99] relocate the saliency prior into self-supervised training but do not remove the photometric problem and monocular endoscopic SLAM with dense depth refinement [100] still requires explicit outlier handling for specularities and tissue deformation. The clinical payoff of endoscopic vision has been established by randomized trials rather than benchmarks, in contrast to most of the other domains surveyed here. Real-time computer-aided detection raises adenoma detection rates in colonoscopy [101] and real-time coverage monitoring reduces blind spots in upper endoscopy [102] the latter is a completeness prior over the observed anatomy, formally analogous to angular coverage in tomography. Frame-level quality assessment coupled with

restoration [103] provides the missing link between the photometric models above and these clinical endpoints, by deciding which frames are informative at all. Our work on automatic feedback in endoscopy [24] sits in this space, closing the loop from image analysis back to the operator during the procedure, where a quality signal can still change what is acquired the one place in medical imaging where the prior can be replaced by a better measurement. Reproducible progress here depends on open benchmarks such as Kvasir-SEG [104], which evaluate segmentation rather than photometric fidelity, a gap that remains.

Gigapixel Whole-Slide Imaging in Computational Pathology

Digital pathology appears at first to break the framework of this review, because nothing is missing. A whole-slide image at diagnostic magnification contains on the order of 10^9 to 10^{10} pixels, decomposable into 10^4 to 10^5 informative tiles and the entire tissue section is present. The framework survives because the operative constraint is the same: No model consumes a gigapixel field directly, so every method imposes a reduction prior a statement about which parts of the data matter and how the remainder may be discarded and the accuracy of that statement determines performance exactly as the accuracy of a smoothness prior determines a reconstruction. Incompleteness is reintroduced deliberately, by the algorithm and the design question is where to reintroduce it. The dominant answer decomposes the problem into tiling, tile encoding and aggregation. Aggregation is where the prior is most explicit. Multiple-instance learning treats a slide as a bag of tiles with a single label and attention-based deep MIL [105] supplies a permutation-invariant gated pooling whose attention weights are precisely a learned prior over which regions carry the label. CLAM [106] added clustering constraints and instance-level supervision derived from the attention distribution, becoming the reference implementation for data-efficient weak supervision and Campanella et al. [107] established at a scale of 44,732 slides that slide-level labels alone suffice to reach clinical decision thresholds the empirical argument that made weak supervision the field's default. The subsequent literature has largely consisted of relaxing MIL's independence assumption: TransMIL [108] replaces independent instances with self-attention over tiles, imposing a correlated spatial prior; DSMIL [109] adds a critical-instance stream and an explicit multi-resolution pyramid; HIPT [110] encodes the pyramid in the architecture itself through a three-level hierarchy from cell to slide. Tile encoding has been transformed by self-supervised

pretraining and this is the fifth stage of the prior's migration in its clearest form.

In-domain contrastive pretraining [111] first displaced ImageNet initialization; the current generation trains vision transformers on hundreds of millions to billions of pathology tiles UNI on more than 100 million tiles from 100,000 slides [13], Virchow on approximately 1.5 million clinical slides with demonstrated scaling behaviour and rare-cancer generalization [112] and Prov-GigaPath on 1.3 billion tiles with dilated attention providing slide-level context across the full tile sequence [14]. The practical consequence is that these embeddings function as a task-agnostic dictionary in which downstream clinical heads become nearly linear probes the same role that a learned sparsifying dictionary played in Section 3, at four orders of magnitude greater scale. Compression addresses the reduction problem information-theoretically rather than through attention and is the strand our own recent work has pursued. Neural image compression for gigapixel histopathology [113] established the principle: Replace the image pyramid with a learned latent map compact enough for end-to-end processing. The open question is how much rate reduction a downstream diagnostic model tolerates before performance degrades and the answer is content- and task-dependent, which argues for adaptivity rather than a fixed codec.

Our adaptive compression framework for gigapixel whole-slide images [27,114] allocates rate according to diagnostic relevance rather than uniformly, so that regions the model will attend to are preserved at high fidelity while diagnostically inert stroma and background are aggressively reduced. Framed in this review's terms, compression and attention are the same prior applied at different points in the pipeline one discards before encoding, the other after and combining them is more effective than either alone because the redundancy they exploit is different. The field's central weakness is covariate shift. Stain appearance varies with reagent lot, protocol, scanner and institution and models trained at one site degrade at another in ways that internal validation cannot reveal. Classical stain-vector deconvolution [115] imposes a hand-designed colour prior and remains the standard baseline; controlled multi-centre comparison established that heavy colour augmentation generally outperforms normalization [116], which is the same lesson the reconstruction literature learned about implicit versus explicit priors. The most severe instance is the intraoperative gap: Frozen sections exhibit freezing artifact, folds and inferior morphology relative to the formalin-fixed paraffin-embedded slides on which models are trained, yet frozen-section diagnosis is where turnaround time makes automated support most valuable. Generative style translation

between the domains [117] is one response. Ours is to align features rather than appearance: Joint feature alignment combined with self-knowledge distillation for tumour detection in frozen-section histopathology [28] matches representations across the domain gap while distilling the model's own predictions to stabilize training on the noisier domain a strategy related to self-distillation within MIL [118], where bag-level supervision is progressively refined into instance-level signal. The most consequential application of these representations is the prediction of molecular state from morphology. Coudray et al. [119] first demonstrated that tile-level networks recover driver-mutation status from H&E in lung cancer; Kather et al. [120] established microsatellite instability prediction in gastrointestinal cancer as the canonical deep-learning biomarker and a subsequent pan-cancer screen [121] mapped which alterations are morphologically predictable at all, with per-cohort ceilings. For gastric cancer the label space is supplied by the four molecular subtypes defined by The Cancer Genome Atlas Research Network [122] Epstein-Barr virus positive, microsatellite unstable, genomically stable and chromosomally unstable which differ in prognosis and treatment implication but require molecular assays to assign. Our weakly supervised TransMIL-attention framework for predicting these subtypes from H&E whole-slide images [29] is an attempt to close that gap and it inherits both the promise and the interpretive difficulty of the approach: Attention maps localize the evidence the model used, but attention is saliency, not proof and the spatial support of a morphology-genotype correlation is rarely validated against orthogonal ground truth. The frontier is language. Vision-language pretraining on paired image-text data PLIP from social-media-sourced pairs [123], CONCH from 1.17 million curated image-caption pairs [124] make text simultaneously a supervisory signal and an output modality and instruction-tuned multimodal assistants [125] generate narrative interpretation rather than scalar labels. This shifts the evaluation problem qualitatively: A subtype prediction is right or wrong, whereas a generated report can be fluent, plausible and clinically misleading. Community benchmarking is beginning to address this and our participation in the pathologist reasoning-guided report generation challenge [126] reflects the specific view that report generation should be evaluated against the diagnostic reasoning a pathologist would apply, not against text-similarity metrics that reward fluent recapitulation. Progress in this domain has always depended on shared benchmarks CAMELYON16 established slide-level algorithm-versus-pathologist comparison [?] and the PANDA challenge quantified generalization loss under scanner and laboratory

shift across countries at a scale no single institution can replicate [127] and report generation will need an equivalent.

Deployment, Infrastructure and Evaluation

Priors determine what a method can recover; infrastructure determines whether it is used. Gi-gapixel inference and real-time interventional guidance impose opposite constraints throughput versus latency and both are architectural rather than algorithmic problems. Service de-composition with independent scaling [128] is the standard response for the throughput case, allowing tiling, encoding, aggregation and reporting to scale separately; our microservice-based architecture for artificial-intelligence Internet-of-Things systems [129] applies this pattern to distributed inference and the same decomposition maps directly onto the whole-slide pipeline. For the latency case, on-premises or on-device inference is often mandatory and the trade-offs among latency, bandwidth and privacy in edge medical computing have been systematically surveyed [130]. Both cases inherit the integrity assumptions of distributed sensing discussed in Section 2 [47, 48], which are rarely stated explicitly in clinical AI papers and which a regulatory submission will eventually require. Data governance is the binding constraint on the learned priors this review has traced, since foundation-scale pretraining requires volumes no single institution holds. Federated learning [131] is the standard route, exchanging model updates rather than patient data and its limitations statistical heterogeneity across sites, update-based leakage and the difficulty of auditing a model no participant fully observes are the practical reason multi-institutional pathology models remain rarer than single-institution ones. Evaluation deserves separate emphasis because it is the field's most systematic weakness. Reconstruction is still largely assessed by PSNR and SSIM [62], which are known not to track diagnostic detectability; the task-based protocols attached to the AAPM challenge [74] are the exception rather than the rule. Pathology is assessed by AUROC on internal splits far more often than on external multi-centre cohorts and PANDA [127] exists precisely because the gap between the two is large. Endoscopy is the outlier and the model to follow: Its central claims rest on randomized trials with clinical endpoints [101,102].

Open Challenges and Future Directions

Hallucination in the null space

This is the defining unsolved problem of learned priors and Frikel and Quinto [67] explain why it is not a bug: In limited-angle geometry, structures outside the acquired

wavefront range are undetermined, so a generative prior that renders them is fabricating anatomy that is by construction consistent with the measurement. Diffusion and untrained priors are more expressive than total variation and therefore fabricate more convincingly. The needed development is not better priors but detection a principled way to mark which parts of a reconstruction are data-determined and which are prior-determined. The microlocal analysis exists; propagating it through a learned reconstruction does not.

Calibrated uncertainty

Related but distinct: Clinical use requires per-voxel or per-region confidence that is calibrated, not merely ordinal. Attention weights in MIL are routinely read as evidence localization despite being trained only to improve a bag-level loss and no current method in either reconstruction or pathology provides intervals a clinician can act on.

Distribution shift

Every domain in this review reports the same failure: scanner, protocol, dose, institution and population shift degrade performance more than architectural differences improve it. Augmentation-as-prior [116], feature alignment and self-distillation [28] and generative domain translation [117] are partial answers; none provides a guarantee and the intraoperative frozen-section gap remains the hardest instance.

Task-based evaluation

The metric problem compounds all of the above. A reconstruction with better SSIM may be worse for lesion detection; a report with higher text-similarity may be diagnostically wrong. Evaluation must be reformulated around clinical tasks, which is expensive and therefore rare the reason our own recent effort has gone toward reasoning-guided rather than similarity-based assessment of generated pathology reports [126].

Toward unified priors

The preceding sections also suggest a prospective research direction rather than a purely retrospective account. If the prior is the invariant across sparse-view reconstruction, few-view interventional geometry, photometric correction and gigapixel reduction, then a genuinely shared representation should be possible a learned prior over anatomy and tissue appearance that serves reconstruction, registration and diagnosis rather than being retrained per task. Pathology foundation models [13,14] demonstrate this within one modality. Whether it extends across the acquisition-intervention-diagnosis pipeline remains an important open question for the field.

Conclusion

The development reviewed here is better understood as a continuous progression than as a series of discrete replacements. Quadratic smoothness constraints were succeeded by total variation, total variation by patch self-similarity and learned dictionaries and these in turn by discriminative networks, unrolled iterative schemes, untrained architectures, score-based generative models and most recently foundation-model embeddings combined with attention distributions over gigapixel fields.

Throughout this sequence the mathematical function of the prior has remained the same: to select, from the set of solutions consistent with the measurement, the one that is clinically correct.

Viewed in this way, the outstanding problems of the field can be stated more precisely. They are not primarily problems of expressiveness, since contemporary priors are highly expressive, but rather of accountability namely, determining which components of an output are supported by the measurement and which are supplied by the prior and reporting that distinction with calibrated confidence. Addressing this requirement is likely to be the principal task for the next stage of the field.

Conflicts of Interest

The author declares no competing interests.

Data Availability

No new data were generated for this review.

References

1. Tikhonov AN (1963) Solution of incorrectly formulated problems and the regularization method. Soviet Math Dokl 4: 1035-1038. [Google Scholar]
2. Rudin LI, Osher S, Fatemi E (1992) Nonlinear total variation based noise removal algorithms. Phys D: Nonlinear Phenom 60: 259-268. [Crossref], [Google Scholar]
3. Mumford D, Shah J (1989) Optimal approximations by piecewise smooth functions and associated variational problems. Commun Pure Appl Math 42: 577-685. [Crossref], [Google Scholar]
4. Perona P, Malik J (1990) Scale-space and edge detection using anisotropic diffusion. IEEE Trans Pattern Anal Mach Intell 12: 629-639. [Crossref], [Google Scholar]
5. Buades A, Coll B, Morel JM (2005) A non-local algorithm for image denoising. IEEE 2: 60-65. [Crossref], [Google Scholar]
6. Dabov K, Foi A, Katkovnik V, Egiazarian K (2007) Image denoising by sparse 3-D transform-domain collaborative filtering. IEEE Trans Image Process 16: 2080-2095. [Crossref], [Google Scholar]
7. Elad M, Aharon M (2006) Image denoising *via* sparse and redundant representations over learned dictionaries. IEEE Trans Image Process 15: 3736-3745. [Crossref], [Google Scholar], [Indexed]
8. Zhang K, Zuo W, Chen Y, Meng D, Zhang L (2017) Beyond a Gaussian denoiser: Residual learning of deep CNN for image denoising. IEEE Trans Image Process 26: 3142-3155. [Crossref], [Google Scholar], [Indexed]
9. Sun J, Li H, Xu Z (2016) Deep ADMM-Net for compressive sensing MRI. [Google Scholar]
10. Chen H, Zhang Y, Chen Y, Zhang J, Zhang W, et al. (2018) LEARN: Learned experts' assessment-based reconstruction network for sparse-data CT. IEEE Trans Med Imaging 37: 1333-1347. [Crossref], [Google Scholar], [Indexed]
11. Baguer DO, Leuschner J, Schmidt M (2020) Computed tomography reconstruction using deep image prior and learned reconstruction methods. Inverse Probl 36: 094004. [Crossref], [Google Scholar]
12. Song Y, Shen L, Xing L, Ermon S (2022) Solving inverse problems in medical imaging with score-based generative models. arXiv. [Crossref], [Google Scholar]
13. Chen RJ, Ding T, Lu MY, Williamson DFK, Jaume G, et al. (2024) Towards a general-purpose foundation model for computational pathology. Nat Med 30: 850-862. [Crossref], [Google Scholar], [Indexed]
14. Xu H, Usuyama N, Bagga J, Zhang S, Rao R, et al. (2024) A whole-slide foundation model for digital pathology from real-world data. Nature 630: 181-188. [Crossref], [Google Scholar], [Indexed]
15. Bappy DM, Jeon I (2013) Total variation denoising using X-ray images. 3861-3863. [Google Scholar]
16. Bappy DM, Jeon I (2016) Combination of hybrid median filter and total variation minimisation for medical X-ray image restoration. IET Image Process 10: 261-271. [Crossref], [Google Scholar]
17. Bappy DM, Jeon I (2015) Few view X-ray CT image reconstruction. Proc KSME Annu Meet 3123-3124. [Google Scholar]
18. Bappy DM, Jeon I (2017) Modified simultaneous iterative reconstruction technique for fast, high-quality CT reconstruction. IET Image Process 11: 701-708. [Crossref], [Google Scholar]
19. Bappy DM, Jeon I (2013) X-ray image interpolation for high quality CT reconstruction.
20. Bappy DM, Jeon I (2019) High-quality X-ray computed tomography reconstruction using projected and interpolated images. IET Image Process 13: 1074-1080. [Crossref], [Google Scholar]
21. Nguyen PB, Kang B, Bappy DM, Choi E, Park S, et al. (2018) Real-time microrobot posture recognition *via* biplane X-ray imaging system for external electromagnetic actuation. Int J

- Comput Assist Radiol Surg 13: 1843-1852. [Crossref], [Google Scholar], [Indexed]
22. Bappy DM, Hong A, Choi E, Park JO, Kim CS (2021) Automated three-dimensional vessel reconstruction based on deep segmentation and bi-plane angio-graphic projections. *Comput Med Imaging Graph* 92: 101956. [Crossref], [Google Scholar]
 23. Bappy DM, Kang D, Lee J, Baek H (2024) Prior-based enhanced ASD-POCS for artifact suppression and structural preservation in sparse-view CBCT. *ACM Digital Library* 496-505. [Crossref], [Google Scholar]
 24. DM Bappy, D Kang, J Lee, Y Lee, M Koo, H Baek (2025) Advanced endoscopy imaging with automatic feedback. Springer 15311: 62-78. [Crossref], [Google Scholar]
 25. Bappy DM, Kang D, Lee J, Lee Y, Baek H (2025) Deep prior based limited-angle tomography. Springer 15311: 79-95. [Crossref], [Google Scholar]
 26. Bappy DM, Kang D, Lee J, Lee Y, Koo M, et al. (2025) Specular region detection and covariant feature extraction. Springer 15312: 170-186. [Crossref], [Google Scholar]
 27. Lee J, Takemaru L, Bappy DM, Jeong YS, Jeong WK, et al. (2026) Adaptive compression framework for giga-pixel whole slide images. *Nat Commun* 17: 207. [Crossref], [Google Scholar], [Indexed]
 28. Kim E, Lee YJ, Lee J, Lee H, Bappy DM, et al. (2025) Fus-ing domains for tumor detection in frozen section histopathology using joint fea-ture alignment and self-knowledge distillation. *IEEE* 13: 190745-190753. [Crossref], [Google Scholar]
 29. Jeong Y, Bappy DM, Ahn S, Lee SH (2026) Abstract 1445: Predicting TCGA molecular subtypes of gastric cancer from H&E whole-slide images using a weakly super-vised TransMIL-attention framework. *Cancer Res* 86: 1445. [Crossref], [Google Scholar]
 30. Bertalmio M, Sapiro G, Caselles V, Ballester C (2000) Image inpainting. *ACM Digit Library* 2000: 417-424. [Crossref], [Google Scholar]
 31. Efros AA, Leung TK (1999) Texture synthesis by non-parametric sampling. 1033-1038. [Crossref], [Google Scholar]
 32. Efros AA, Freeman WT (2001) Image quilting for texture synthesis and transfer. *ACM Digit Library* 2001: 341-346. [Crossref], [Google Scholar]
 33. Criminisi A, Pérez P, Toyama K (2004) Region filling and object removal by exemplar-based image inpainting. *IEEE Trans Image Process* 13: 1200-1212. [Crossref], [Google Scholar], [Indexed]
 34. Rahman MH, Yusuf MSU, Bappy DM, Islam MZ, Dey AK (2012) Object duplication in an image using texture synthesis technique. *Int J Eng Tech* 2: 460-462. [Google Scholar]
 35. Pathak D, Krahenbuhl P, Donahue J, Darrell T, Efros AA (2016) Context encoders: Feature learning by inpainting. *IEEE* 2016: 2536-2544. [Crossref], [Google Scholar]
 36. Longuet-Higgins HC (1981) A computer algorithm for reconstructing a scene from two projections. *Nature* 293: 133-135. [Crossref], [Google Scholar]
 37. Metz CE, Fencil LE (1989) Determination of three-dimensional structure in biplane radiography without prior knowledge of the relationship between the two views: Theory. *Med Phys* 16: 45-51. [Crossref], [Google Scholar], [Indexed]
 38. Tomasi C, Kanade T (1992) Shape and motion from image streams under orthography: A factorization method. *Int J Comput Vis* 9: 137-154. [Crossref], [Google Scholar]
 39. Schonberger JL, Frahm JM (2016) Structure-from-motion revisited. *IEEE* 2016: 4104-4113. [Crossref], [Google Scholar]
 40. DM Bappy, MH Rahman (2012) A study in 3D structure detection implementing forward camera motion. [Google Scholar]
 41. Bappy DM, Jeon I (2015) Motion vector estimation for transparent sequences. *Proce KSME Annu Meet* 171-172. [Google Scholar]
 42. Shannon CE (1948) A mathematical theory of communication. *Bell Syst Tech J* 27: 379-423. [Crossref], [Google Scholar]
 43. Weinstein S, Ebert P (1971) Data transmission by frequency-division multiplexing using the discrete fourier transform. *IEEE Trans Commun Technol* 19: 628-634. [Crossref], [Google Scholar]
 44. Bingham JAC (1990) Multicarrier modulation for data transmission: An idea whose time has come. *IEEE Commun Mag* 28: 5-14. [Crossref], [Google Scholar]
 45. Bappy DM, Dey AK, Saha S, Saha A, Ghosh S (2010) OFDM system analysis for reduction of inter symbol interference using the AWGN channel platform. *Int J Adv Comput Sci Appl* 1: 123-125. [Crossref], [Google Scholar]
 46. Akyildiz IF, Su W, Sankarasubramaniam Y, Cayirci E (2002) Wireless sensor networks: A survey. *Comput Netw* 38: 393-422. [Crossref], [Google Scholar]
 47. Islam MS, Khan RH, Bappy DM (2010) A hierarchical intrusion detection system in wireless sensor networks. *Int J Comput Sci Netw Secur* 10: 21-26. [Google Scholar]
 48. Butun I, Morgera SD, Sankar R (2014) A survey of intrusion detection systems in wireless sensor networks. *IEEE Commun Surv Tutor* 16: 266-282. [Crossref], [Google Scholar]
 49. Geman D, Yang C (1995) Nonlinear image recovery with half-quadratic regularization. *IEEE Trans Image Process* 4: 932-946. [Crossref], [Google Scholar], [Indexed]
 50. Goldstein T, Osher S (2009) The split Bregman method for L1-regularized problems. *SIAM J Imaging Sci* 2: 323-343. [Crossref], [Google Scholar]
 51. Chambolle A, Pock T (2011) A first-order primal-dual algorithm for convex problems with applications to imaging. *J Math Imaging Vis* 40: 120-145. [Crossref], [Google Scholar]
 52. A Beck, M Teboulle (2009) A fast iterative shrinkage-thresholding algorithm for linear inverse problems. *SIAM J Imaging Sci* 2: 183-202. [Crossref], [Google Scholar],
 53. Bappy DM, Jeon I (2015) Medical image restoration using L1-minimization. *Proc KSME Annu Meet* 3-4. [Google Scholar]

54. DM Bappy, Jeon I (2014) Fast medical image restoration using variable half-quadratic penalty parameter. *Proc KSME Annu Meet* 2595-2597. [Google Scholar]
55. Hwang H, Haddad RA (1995) Adaptive median filters: New algorithms and results. *IEEE Trans Image Process* 4: 499-502. [Crossref], [Google Scholar], [Indexed]
56. Chan RH, Ho CW, Nikolova M (2005) Salt-and-pepper noise removal by median-type noise detectors and detail-preserving regularization. *IEEE Trans. Image Process* 14: 1479-1485. [Crossref], [Google Scholar]
57. Liang J, Cao J, Sun G, Zhang K, Gool L Van, et al. (2021) SwinIR: Image restoration using swin transformer. *IEEE/CVF Int Conf Comput Vis Workshops 2021*: 1833-1844. [Crossref], [Google Scholar]
58. SV Venkatakrishnan, CA Bouman, Wohlberg B (2013) Plug-and-play priors for model based reconstruction. *Glob Conf Signal Inf Process 2013*: 945-948. [Crossref], [Google Scholar]
59. Gordon RR, Bender R, Herman GT (1970) Algebraic Reconstruction Techniques (ART) for three-dimensional electron microscopy and X-ray photography. *J Theor Biol* 29: 471-481. [Crossref], [Google Scholar], [Indexed]
60. Gilbert P (1972) Iterative methods for the three-dimensional reconstruction of an object from projections. *J Theor Biol* 36: 105-117. [Crossref], [Google Scholar], [Indexed]
61. Andersen AH, Kak AC (1984) Simultaneous Algebraic Reconstruction Technique (SART): A superior implementation of the ART algorithm. *Ultrason Imaging* 6: 81-94. [Crossref], [Google Scholar], [Indexed]
62. Wang Z, Bovik AC, Sheikh HR, Simoncelli EP (2004) Image quality assessment: From error visibility to structural similarity. *IEEE Trans Image Process* 13: 600-612. [Crossref], [Google Scholar], [Indexed]
63. Shepp LA, Vardi Y (1982) Maximum likelihood reconstruction for emission tomography. *IEEE Trans Med Imaging* 1: 113-122. [Crossref], [Google Scholar], [Indexed]
64. Sidky EY, Kao CM, Pan X (2006) Accurate image reconstruction from few-views and limited-angle data in divergent-beam CT. *J X-Ray Sci Technol* 14: 119-139. [Crossref], [Google Scholar]
65. EY Sidky, X Pan (2008) Image reconstruction in circular cone-beam computed tomography by constrained, total-variation minimization. *Phys Med Biol* 53: 4777-4807. [Crossref], [Google Scholar], [Indexed]
66. Chen GH, Tang J, Leng S (2008) Prior Image Constrained Compressed Sensing (PICCS): A method to accurately reconstruct dynamic CT images from highly undersampled projection data sets. *Med Phys* 35: 660-663. [Crossref], [Google Scholar], [Indexed]
67. Friel J, Quinto ET (2013) Characterization and reduction of artifacts in limited angle tomography. *Inverse Probl* 29: 125007. [Crossref], [Google Scholar]
68. Bappy DM, Jeon I (2016) A modified CGTV algorithm for X-ray CT reconstruction. *37*: 37. [Google Scholar]
69. Lee H, Lee J, Kim H, Cho B, Cho S (2019) Deep-neural-network-based sinogram synthesis for sparse-view CT image reconstruction. *IEEE Trans Radiat Plasma Med Sci* 3: 109-119. [Crossref], [Google Scholar]
70. Jin KH, McCann MT, Froustey E, Unser M (2017) Deep convolutional neural network for inverse problems in imaging. *IEEE Trans Image Process* 26: 4509-4522. [Crossref], [Google Scholar], [Indexed]
71. Chen H, Zhang Y, Kalra MK, Lin F, Chen Y, et al. (2017) Low-dose CT with a residual encoder-decoder convolutional neural network. *IEEE Trans Med Imaging* 36: 2524-2535. [Crossref], [Google Scholar], [Indexed]
72. Yang Q, Yan P, Zhang Y, Yu H, Shi Y, et al. (2018) Low-dose CT image denoising using a generative adversarial network with Wasserstein distance and perceptual loss. *IEEE Trans Med Imaging* 37: 1348-1357. [Crossref], [Google Scholar], [Indexed]
73. Zhu B, Liu JZ, Cauley SF, Rosen BR, Rosen MS (2018) Image reconstruction by domain-transform manifold learning. *Nature* 555: 487-492. [Crossref], [Google Scholar], [Indexed]
74. McCollough CH, Bartley AC, Carter RE, Chen B, Drees TA, et al. (2017) Low-dose CT for the detection and classification of metastatic liver lesions: Results of the 2016 Low Dose CT Grand Challenge. *Medi Phys* 44: e339-e352. [Crossref], [Google Scholar], [Indexed]
75. Balter S, Hopewell JW, Miller DL, Wagner LK, Zelefsky MJ (2010) Fluoroscopically guided interventional procedures: A review of radiation effects on patients' skin and hair. *Radiology* 254: 326-341. [Crossref], [Google Scholar], [Indexed]
76. Crummy AB, Strother CM, Sackett JF, Ergun DL, Shaw CG, et al. (1980) Computerized fluoroscopy: Digital subtraction for intravenous angiocardiology and arteriography. *AJR Am J Roentgenol* 135: 1131-1140. [Crossref], [Google Scholar], [Indexed]
77. Chen SJ, Carroll JD (2000) 3-D reconstruction of coronary arterial tree to optimize angiographic visualization. *IEEE Trans Med Imaging* 19: 318-336. [Crossref], [Google Scholar], [Indexed]
78. Blondel C, Malandain G, Vaillant R, Ayache N (2006) Reconstruction of coronary arteries from a single rotational X-ray projection sequence. *IEEE Trans Med Imaging* 25: 653-663. [Crossref], [Google Scholar], [Indexed]
79. Çimen S, Gooya A, Grass M, Frangi AF (2016) Reconstruction of coronary arteries from X-ray angiography: A review. *Med Image Anal* 32: 46-68. [Crossref], [Google Scholar], [Indexed]
80. Frangi AF, Niessen WJ, Vincken KL, Viergever MA (1998) Multiscale vessel enhancement filtering. *Springer* 1496: 130-137. [Crossref], [Google Scholar]
81. Ronneberger O, Fischer P, Brox T (2015) U-Net: Convolutional networks for biomedical image segmentation. *Springer* 9351: 234-241. [Crossref], [Google Scholar]
82. Staal J, Abramoff MD, Niemeijer M, Viergever MA, Ginneken B Van (2004) Ridge-based vessel segmentation in color images of

- the retina. *IEEE Trans Med Imaging* 23: 501-509. [Crossref], [Google Scholar], [Indexed]
83. Markelj P, Tomaževič D, Likar B, Pernuš F (2012) A review of 3D/2D registration methods for image-guided interventions. *Med Image Anal* 16: 642-661. [Crossref], [Google Scholar], [Indexed]
84. Ambrosini P, Ruijters D, Niessen WJ, Moelker A, Walsum T van (2017) Fully automatic and real-time catheter segmentation in X-ray fluoroscopy. *Springer 10434*: 577-585. [Crossref], [Google Scholar]
85. Wang Y, Banerjee A, Grau V (2024) NeCA: 3d coronary artery tree reconstruction from two 2d projections *via* neural implicit representation. *Bioengi* 11: 1227. [Crossref], [Google Scholar], [Indexed]
86. Nelson BJ, Kaliakatsos IK, Abbott JJ (2010) Microrobots for minimally invasive medicine. *Annu Rev Biomed Eng* 12: 55-85. [Crossref], [Google Scholar], [Indexed]
87. Kummer MP, Abbott JJ, Kratochvil BE, Borer R, Sengul A (2010) OctoMag: An electromagnetic system for 5-DOF wireless micromanipulation. *IEEE Trans Robot* 26: 1006-1017. [Crossref], [Google Scholar]
88. Aziz A, Pane S, Iacovacci V, Koukourakis N, Czarske J, et al. (2020) Medical imaging of microrobots: Toward *in vivo* applications. *ACS Nano* 14: 10865-10893. [Crossref], [Google Scholar], [Indexed]
89. Ligtenberg LJW, Rabou NCA, Goulas C, Duinmeijer WC, Halfwerk FR, et al. (2024) *Ex vivo* validation of magnetically actuated intravascular untethered robots in a clinical setting. *Commun Eng* 3: 68. [Crossref], [Google Scholar], [Indexed]
90. Shafer SA (1985) Using color to separate reflection components. *Color Res Appl* 10: 210-218. [Crossref], [Google Scholar]
91. Tan RT, Ikeuchi K (2005) Separating reflection components of textured surfaces using a single image. *IEEE Trans Pattern Anal Mach Intell* 27: 178-193. [Crossref], [Google Scholar], [Indexed]
92. Arnold M, Ghosh A, Ameling S, Lacey G (2010) Automatic segmentation and inpainting of specular highlights for endoscopic imaging. *J Image Video Proc* 2010: 814319, [Crossref], [Google Scholar]
93. Daher R, Vasconcelos F, Stoyanov D (2023) A temporal learning approach to inpainting endoscopic specularities and its effect on image correspondence. *Med Image Anal* 90: 102994. [Crossref], [Google Scholar], [Indexed]
94. Lowe DG (2004) Distinctive image features from scale-invariant keypoints. *Int J Comput Vis* 60: 91-110. [Crossref], [Google Scholar]
95. Bay H, Ess A, Tuytelaars T, Gool L Van (2008) Speeded-Up Robust Features (SURF). *Comput Vis Image Underst* 110: 346-359. [Crossref], [Google Scholar]
96. Rublee E, Rabaud V, Konolige K, Bradski G (2011) ORB: An efficient alternative to SIFT or SURF. *Proc IEEE Int Conf Comput Vis* 2011: 2564-2571. [Crossref], [Google Scholar]
97. Mur-Artal R, Montiel JMM, Tardos JD (2015) ORB-SLAM: A versatile and accurate monocular SLAM system. *IEEE Trans Robot* 31: 1147-1163. [Crossref], [Google Scholar]
98. Mikolajczyk K, Tuytelaars T, Schmid C, Zisserman A, Matas J, et al. (2005) A comparison of affine region detectors. *Int J Comput Vis* 65: 43-72. [Crossref], [Google Scholar]
99. DeTone D, Malisiewicz T, Rabinovich A (2018) SuperPoint: Self-supervised interest point detection and description. *IEEE* 224-236. [Crossref], [Google Scholar]
100. Mahmoud N, Collins T, Hostettler A, Soler L, Doignon C (2019) Live tracking and dense reconstruction for handheld monocular endoscopy. *IEEE Trans Med Imaging* 38: 79-89. [Crossref], [Google Scholar], [Indexed]
101. Repici A, Badalamenti M, Maselli R, Correale L, Radaelli F, et al. (2020) Efficacy of real-time computer-aided detection of colorectal neoplasia in a randomized trial. *Gastroenterology* 159: 512-520.e7. [Crossref], [Google Scholar], [Indexed]
102. Wu L, Zhang J, Zhou W, An P, Shen L, et al. (2019) Randomised controlled trial of WISENSE, a real-time quality improving system for monitoring blind spots during esophagogastroduodenoscopy. *Gut* 68: 2161-2169. [Crossref], [Google Scholar], [Indexed]
103. Ali S, Zhou F, Bailey A, Braden B, East JE, et al. (2021) A deep learning framework for quality assessment and restoration in video endoscopy. *Med Image Anal* 68: 101900. [Crossref], [Google Scholar], [Indexed]
104. Jha D, Smedsrud PH, Riegler MA, Halvorsen P, Lange T De, et al. (2020) Kvasir-SEG: A segmented polyp dataset. *Springer* 11962: 451-462. [Crossref], [Google Scholar]
105. Ilse M, Tomczak J, Welling M (2018) Attention-based deep multiple instance learning. *Proc Mach Learn Res* 80: 2127-2136. [Google Scholar]
106. Lu MY, Williamson DFK, Chen TY, Chen RJ, Barbieri M, et al. (2021) Data-efficient and weakly supervised computational pathology on whole-slide images. *Nat Biomed Eng* 5: 555-570. [Crossref], [Google Scholar], [Indexed]
107. Campanella G, Hanna MG, Geneslaw L, Miraflor A, Silva VWK, et al. (2019) Clinical-grade computational pathology using weakly supervised deep learning on whole slide images. *Nat Med* 25: 1301-1309. [Crossref], [Google Scholar], [Indexed]
108. Shao Z, Bian H, Chen Y, Wang Y, Zhang J, et al. (2021) TransMIL: Transformer based correlated multiple instance learning for whole slide image classification. *NeurIPS* 34: 2136-2147. [Google Scholar]
109. Li B, Li Y, Eliceiri KW (2021) Dual-stream multiple instance learning network for whole slide image classification with self-supervised contrastive learning. *Conf Comput Vis Pattern Recognit Workshops* 2021: 14318-14328. [Crossref], [Google Scholar], [Indexed]
110. Chen RJ, Chen C, Li Y, Chen TY, Trister AD, et al. (2022) Scaling vision transformers to gigapixel images *via* hierarchical self-supervised learning. *IEEE Comput Soc Conf Comput Vis*

- Pattern Recognit 2022: 16144-16155. [Crossref], [Google Scholar]
111. Wang XX, Yang S, Zhang J, Wang M, Zhang J, et al. (2022) Transformer-based unsupervised contrastive learning for histopathological image classification. *Med Image Anal* 81: 102559. [Crossref], [Google Scholar], [Indexed]
 112. Vorontsov E, Bozkurt A, Casson A, Shaikovski G, Zelechowski M, et al. (2024) A foundation model for clinical-grade computational pathology and rare cancers detection. *Nat Med* 30: 2924-2935. [Crossref], [Google Scholar], [Indexed]
 113. Tellez D, Litjens G, Van der Laak J, Ciompi F (2021) Neural image compression for gigapixel histopathology image analysis. *IEEE Trans Pattern Anal Mach Intell* 43: 567-578. [Crossref], [Google Scholar], [Indexed]
 114. Ahn S, Lee J, Bappy D, Jeong WK, Lee SH (2024) Adaptive compression framework for giga-pixel whole slide images. *Res Sq* [Crossref], [Google Scholar]
 115. Macenko M, Niethammer M, Marron JS, Borland D, Woosley JT, et al. (2009) A method for normalizing histology slides for quantitative analysis. *IEEE* 1107-1110. [Crossref], [Google Scholar]
 116. Tellez D, Litjens G, Bándi P, Bulten W, Bokhorst JM, et al. (2019) Quantifying the effects of data augmentation and stain color normalization in convolutional neural networks for computational pathology. *Med Image Anal* 58: 101544. [Crossref], [Google Scholar], [Indexed]
 117. Ozyoruk KB, Can S, Darbaz B, Başak K, Demir D, et al. (2022) A deep-learning model for transforming the style of tissue images from cryosectioned to formalin-fixed and paraffin-embedded. *Nat Biomed Eng* 6: 1407-1419. [Crossref], [Google Scholar], [Indexed]
 118. Wu S, Qiu Y, Nearchou IP, Prost S, Fallowfield JA, et al. (2025) LadderMIL: Multiple instance learning with coarse-to-fine self-distillation. *IEEE J Biomed Health Inform* 1-15. [Crossref], [Google Scholar], [Indexed]
 119. Coudray N, Ocampo PS, Sakellaropoulos T, Narula N, Snuderl M (2018) Classification and mutation prediction from non-small cell lung cancer histopathology images using deep learning. *Nat Med* 24: 1559-1567. [Crossref], [Google Scholar], [Indexed]
 120. Kather JN, Pearson AT, Halama N, Jäger D, Krause J, et al. (2019) Deep learning can predict microsatellite instability directly from histology in gastrointestinal cancer. *Nat Med* 25: 1054-1056. [Crossref], [Google Scholar], [Indexed]
 121. Kather JN, Heij LR, Grabsch HI, Loeffler C, Echle A, et al. (2020) Pan-cancer image-based detection of clinically actionable genetic alterations. *Nat Cancer* 1: 789-799. [Crossref], [Google Scholar], [Indexed]
 122. (2014) Comprehensive molecular characterization of gastric adenocarcinoma. *Nature* 513: 202-209. [Crossref], [Google Scholar], [Indexed]
 123. Huang Z, Bianchi F, Yuksekgonul M, Montine TJ, Zou J (2023) A visual-language foundation model for pathology image analysis using medical Twitter. *Nat Med* 29: 2307-2316. [Crossref], [Google Scholar], [Indexed]
 124. Lu MY, Chen B, Williamson DFK, Chen RJ, Liang I (2024) A visual-language foundation model for computational pathology. *Nat Med* 30: 863-874. [Crossref], [Google Scholar]
 125. Lu MY, Chen B, Williamson DFK, Chen RJ, Zhao M, et al. (2024) A multimodal generative AI copilot for human pathology. *Nature* 634: 466-473. [Crossref], [Google Scholar], [Indexed]
 126. (2026) Pathologist REasoning-Guided REport generation challenge.
 127. Bulten W, Kartasalo K, Chen PHC, Ström P, Pinckaers H, et al. (2022) Artificial intelligence for diagnosis and gleason grading of prostate cancer: The PANDA challenge. *Nat Med* 28: 154-163. [Crossref], [Google Scholar], [Indexed]
 128. Dragoni N, Giallorenzo S, Lafuente AL, Mazzara M, Montesi F, et al. (2017) Microservices: Yesterday, today and tomorrow. *Springer* 195-216. [Crossref], [Google Scholar]
 129. Kim Y, BAPPY DM, Kim J (2021) Artificial intelligence IoT architecture based on microservices. *J Cont Comp* 3: 289-295. [Google Scholar]
 130. Amin SU, Hossain MS (2021) Edge intelligence and Internet of things in healthcare: A survey. *IEEE* 9: 45-59. [Crossref], [Google Scholar]
 131. Rieke N, Hancox J, Li W, Milletari F, Roth HR, et al. (2020) The future of digital health with federated learning. *NPJ Digit Med* 3: 119. [Crossref], [Google Scholar], [Indexed]