Medical Image Compression in the Era of Clinical AI:

The Fidelity–Utility Dilemma and Information Preservation

İnci Zaim Gökbay • Bekir Sıddık Binboğa Yarman

Compression preserves pixels, not decisions

ÖZGÜR PRESS

Medical Image Compression in the Era of Clinical AI: The Fidelity–Utility Dilemma and Information Preservation

Asst. Prof. Dr. İnci Zaim Gökbay
Prof. Dr. Bekir Sıddık Binboğa Yarman

ÖZGÜR PRESS

Published by
Özgür Yayın-Dağıtım Co. Ltd.
Certificate Number: 45503

15 Temmuz Mah. 148136. Sk. No: 9 Şehitkamil/Gaziantep
+90.850 260 09 97
+90.532 289 82 15
www.ozguryayınlari.com
info@ozguryayinlari.com

Medical Image Compression in the Era of Clinical AI: The Fidelity–Utility Dilemma and Information Preservation

Asst. Prof. Dr. İnci Zaim Gökbay • Prof. Dr. Bekir Sıddık Binboğa Yarman

Language: English
Publication Date: 2026
Cover design by Mehmet Çakır
Cover design and image licensed under CC BY-NC 4.0
Print and digital versions typeset by Çizgi Medya Co. Ltd.

ISBN (PDF): 978-625-8998-61-0

DOI: https://doi.org/10.58830/ozgur.pub1277

This work is licensed under the Creative Commons Attribution-NonCommercial 4.0 International (CC BY-NC 4.0). To view a copy of this license, visit https://creativecommons.org/licenses/by-nc/4.0/ This license allows for copying any part of the work for personal use, not commercial use, providing author attribution is clearly stated.

Suggested citation:
Zaim Gökbay, İ., Yarman, B. S. B. (2026). *Medical Image Compression in the Era of Clinical AI: The Fidelity–Utility Dilemma and Information Preservation.* Özgür Publications. DOI: https://doi.org/10.58830/ozgur.pub1277. License: CC-BY-NC 4.0

The full text of this book has been peer-reviewed to ensure high academic standards. For full review policies, see https://www.ozguryayinlari.com/

ÖZGÜR PRESS

Dedicated to the memory of Güneş Cömert, who brought meaning and beauty to the lives of both authors

Preface

This monograph was born from a fundamental scientific curiosity—a relentless, meticulous drive to decompose an accepted engineering standard down to its most basic epistemological roots. It is the culmination of a journey that began as a rigorous dialog between a doctoral candidate and her mentor, questioning whether the mathematical perfection of an algorithm truly aligned with biological and diagnostic reality.

The theoretical groundwork for this inquiry derives from the first author's doctoral dissertation, "Modeling of Biomedical Images via the SYMPES Method" (Gökbay, 2013), completed under the supervision of the second author at Istanbul University. The Classical SYMPES model developed therein was a rigorous instantiation of classical rate–distortion theory, designed to minimize quadratic distortion in intensity space. Under this objective, compression ratios up to 51× were achieved while maintaining PSNR values above 31 dB in moderate-rate regimes. From a pure signal-processing perspective, the equations balanced; the problem appeared solved.

Yet, an analytical mind cannot rest on scalar metrics alone. Discussions with clinicians and a deeper, evidence-based dissection of the results revealed a question that was no longer engineering-based, but epistemological: Was diagnostic information truly preserved?

Classical information theory addresses fidelity using Mean Squared Error (MSE), PSNR, and Structural Similarity (SSIM), implicitly asserting that preserving pixel variance equates to preserving image content. We found that this equivalence holds only within the geometry of intensity space. Clinical decisions—whether rendered by human radiologists or deep neural networks—operate in a structurally distinct realm: representation space, governed by radiomic feature manifolds, morphological topology, and learned embeddings.

Minimization of quadratic distortion does not guarantee stability in these spaces. As we applied Classical SYMPES not merely as a codec, but as an analytically transparent diagnostic instrument, the empirical truth emerged. A compressed CT image may yield PSNR values ≥ 35 dB while inducing measurable topological perturbations or altering discriminative radiomic signatures. In our TCGA-LUAD stress tests, SSIM values as high as 0.991 coexisted with a 9.3% relative decline in oncological staging AUC. Pixel fidelity remained asymptotically stable; representation fidelity drifted.

We term this structural asymmetry the Variance–Information Fallacy: variance is not a sufficient statistic for diagnostic information.

This monograph revisits the foundational assumptions of classical compression in the context of tomographic imaging. The aim is not to introduce a new codec per se, but to fundamentally reformulate the optimization objective—shifting from pixel-wise energy minimization to task-conditioned structural stability in representation space. These results do not invalidate classical distortion minimization; rather, they demonstrate its incompleteness when the ultimate objective is clinical decision stability in the age of artificial intelligence.

Structure of the Book

Part I formalizes CT not as a deterministic camera but as a stochastic, ill-posed measurement and inference system, in which operator-induced variance becomes entangled with anatomical truth.

Part II dissects Shannon's quadratic rate–distortion framework, explaining why classical compression networks naturally fall victim to the Variance–Information Fallacy when applied to medical data.

Part III presents a systematic empirical analysis using the multi-center TCGA-LUAD cohort. It demonstrates that while global metrics like PSNR remain stable, clinically vital structures experience non-monotonic phase transitions and topological collapse under increasing compression.

Part IV generalizes these empirical anomalies into a rigorous mathematical framework. It proves the Fidelity–Topology Decoupling Theorem and introduces the Rate–Structure Paradigm, defining clinical safety through topological invariants and manifold convergence.

Part V advances this framework into the era of Clinical AI. It models compression as a geometric perturbation operator within deep learning pipelines, demonstrating how subtle artifacts cause "manifold drift" in latent

feature spaces, and advocates for a shift toward task-aware, representation-preserving compression.

This monograph is written for researchers in medical image processing, radiological physics, biomedical engineering, and quantitative imaging who share our intolerance for assumptions and demand rigorous, evidence-based methodologies. The required foundations in linear algebra, probability theory, inverse problems, and persistent homology are developed systematically, step by step, throughout the text.

At its core, this work invites the scientific community to reconsider a tacit assumption: that minimizing error in intensity space is equivalent to preserving truth. In medical imaging, truth is not defined by scalar variance alone, but by the unshakeable stability of structure across the clinical decision manifold.

Asst. Prof. Dr. İnci Zaim Gökbay

Prof. Dr. Bekir Sıddık Binboğa Yarman

Istanbul, 2026.

Contents

List of Figures

List of Tables

List of Abbreviations

AAPM	American Association of Physicists in Medicine
AI	Artificial Intelligence
AUC	Area Under the Curve
BPP	Bits Per Pixel
CEB	Classified Energy Bank
CEP	Circular Error Probable
CPB	Classified Pattern Bank
CR	Compression Ratio
CT	Computed Tomography
DCE	Dynamic Contrast-Enhanced
DCT	Discrete Cosine Transform
DICOM	Digital Imaging and Communications in Medicine
FBP	Filtered Backprojection
GLCM	Gray Level Co-occurrence Matrix
HU	Hounsfield Unit
IRR	Information Retention Ratio
JPEG	Joint Photographic Experts Group
MI	Mutual Information
MRI	Magnetic Resonance Imaging
MSE	Mean Squared Error
PACS	Picture Archiving and Communication System
PCA	Principal Component Analysis
PSNR	Peak Signal-to-Noise Ratio
RECIST	Response Evaluation Criteria in Solid Tumors
ROI	Region of Interest
ROI-L	Region of Interest for Lung Tumor
ROI-X	Extended Region of Interest

SNR	Signal-to-Noise Ratio
SSIM	Structural Similarity Index Measure
SVD	Singular Value Decomposition
SYMPES	Systematic Procedure for Predefined Envelope and Signature Sequences (Gökbay, 2013)
TC-SYMPES	Task-Conditioned SYMPES
TCGA	The Cancer Genome Atlas
TCGA-LUAD	TCGA Lung Adenocarcinoma Collection
TCIA	The Cancer Imaging Archive
TV	Total Variation

List of Symbols

x	Original medical image
$\dot{x}$ ($\tilde{x}$)	Compressed image after compression operator
$\hat{x}$	Reconstructed image after decompression
z	Latent feature representation, $z = f\theta(x)$
$\tilde{z}$	Perturbed latent representation under compression
$f\theta$	Parameterized feature extraction function
g	Clinical decision/inference function
$C\lambda$	Compression operator at level λ
θ	Model parameters or projection angle
λ	Compression level / Lagrange multiplier
α	Task loss weighting coefficient
β	Rate weighting coefficient in joint optimization
ε	Tolerance parameter for feature stability
$\mu(r)$	Linear attenuation coefficient at position r
σ^2	Variance of signal or noise
I_0	Incident X-ray intensity
$D(x, \tilde{x})$	Pixel-level distortion measure
R	Bitrate (bits per pixel)
L_task	Task-dependent loss function
$H(X)$	Shannon entropy of source X
$I(X; Y)$	Mutual information between X and Y
R_info	Information retention ratio
D_latent	Average latent displacement under compression
D_class	Class-center deviation in feature space
R_pred	Predictive consistency rate
Δz	Feature perturbation vector, $\Delta z = \tilde{z} - z$
$J\_f(x)$	Jacobian of feature extractor at x

$\nabla g(z)$	Gradient of classifier at z
M_k	Latent manifold associated with class k
$\Delta\beta_k$	Change in k-th Betti number under compression
d_B	Bottleneck distance between persistence diagrams
$Dgm(X)$	Persistence diagram of image X
$\|\cdot\|_1$	L_1 norm (sum of absolute values)
$\|\cdot\|_2$	L_2 norm (Euclidean norm)
$\otimes$	Convolution operator
$\odot$	Hadamard (element-wise) product
C_{safe}	Clinical safety criterion for compression

CHAPTER 1

CT as a Statistical Measurement System

Chapter Thesis

Medical images are statistical reconstructions of photon measurements rather than direct anatomical records.

1.1 The Forward Model and the Emergence of the Inverse Problem

Although computed tomography (CT) is clinically defined as a "cross-sectional imaging" modality, at a physical and mathematical level, it is a measurement–inference system. The CT scanner does not observe anatomical structures directly. What it measures is the intensity of X-rays traversing the tissue upon reaching the detector.

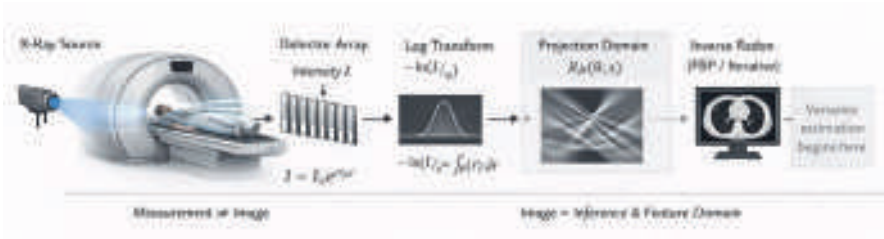

Figure 1.1. Computed tomography forward model. An X-ray source emits photons that traverse the attenuation field $\mu(r)$***, and are recorded by the detector array as transmitted intensity*** I***. The measurement follows the Beer–Lambert law***

$$I(\theta,s) = I_0 \exp\left(-\int_{L(\theta,s)} \mu(r)\, dl\right)$$

This distinction is critical: the scanner records photon attenuation, not anatomy itself. The geometric and physical structure of this forward model is summarized in Figure 1.1. The quantity measured by the detector, $I(\theta, s)$, is the transmitted intensity of a ray propagating at a specific angle θ and position s in space. However, this is not the quantity of clinical interest. Our objective is to recover the spatial attenuation distribution of the tissue, namely the function $\mu(r)$.

Consequently, the forward problem and the inverse problem in CT are explicitly bifurcated:

- **Forward problem:** If the attenuation distribution is known, the detector signal can be computed.
- **Inverse problem:** The attenuation distribution is sought to be recovered from the detector signal.

The mathematical essence of tomography resides in this inversion process.

1.1.1. The Beer–Lambert Law and Heterogeneous Media

For a monoenergetic X-ray beam, the Beer–Lambert law dictates:

$$I = I_0 \exp(-\mu x) \tag{1.1}$$

Where:

I_0 : Initial intensity emitted from the source

I : Intensity reaching the detector

μ : Linear attenuation coefficient

x : Path length traversed by the beam through the medium (Hsieh, 2015; Bushberg et al., 2011).

This expression presupposes a homogeneous medium. Human anatomy, however, is not homogeneous. Bone, soft tissue, air, and contrast agents possess distinct attenuation coefficients. Therefore, a spatially dependent function, $\mu = \mu(r)$, is implicated. Under these conditions, the transmitted intensity assumes the following form:

$$I(\theta, s) = I_0 \exp\left(-\int_{L(\theta,s)} \mu(r)\, dl\right) \tag{1.2}$$

Thus, the quantity measured by the detector is the exponential function of the line integral of the attenuation field.

1.1.2. The Structure of the Inverse Problem

At this juncture, a critical observation must be articulated: The detector does not measure a linear integral. The detector measures an exponentially attenuated intensity. The target function, $\mu(r)$, is embedded within the exponential function in the measurement equation. Consequently, the measurement operator is non-linear. Linear operators hold a privileged position in inverse problem theory. If the measurement were linear, in the form of $\boldsymbol{y} = \mathrm{A}\boldsymbol{\mu}$:

- The spectral properties of the operator could be investigated,
- The uniqueness of the solution could be analyzed,
- Noise amplification could be quantified (Deans, 2007; Natterer, 2001).

However, the measurement in the Beer–Lambert formulation is structured as $I = I_0 e^{-\hat{A}\mu}$. This architecture is non-linear; the unknown field is enclosed within the exponential. This condition severely complicates the direct analysis of the inverse problem.

1.1.3. Logarithm and Linearization

Precisely at this point, a critical transformation is executed. The measurement is normalized, and its logarithm is taken:

$$-\ln\left(\frac{I(\theta,s)}{I_0}\right) = \int_{L(\theta,s)} \mu(r)\, dl \tag{1.3}$$

This operation bears significance on two levels:

1. **Physical level:** The detector measures intensity.
2. **Mathematical level:** We aim to obtain the line integrals of the attenuation field.

The logarithmic operation linearizes the exponential relationship. The measurement is now a linear function of the unknown target. By definition, the expression:

$$p(\theta,s) = -\ln\left(\frac{I(\theta,s)}{I_0}\right) \tag{1.4}$$

constitutes the projection data of CT. The projection data represents the total attenuation along the ray at a specific angle and position. Therefore, the

logarithmic step is not merely an algebraic convenience; it is a mandatory procedure that transforms the inverse problem into a linear operator format. Without this step, tomography cannot be formulated as a linear integral equation.

1.1.4. The Emergence of the Radon Transform

The resulting expression:

$$p(\theta, s) = \int_{L(\theta,s)} \mu(r)\, dl \tag{1.5}$$

is the definition of the two-dimensional Radon transform. Radon (1917) demonstrated that, under sufficient smoothness and compact support conditions, a function can be uniquely recovered from its integrals over all lines. Because the anatomical attenuation field is continuous and defined over a bounded domain, this finding constitutes the mathematical foundation of tomographic imaging (Feeman, 2010). However, this result is an existence theorem. It guarantees recoverability but does not provide a practical inversion algorithm.

In the tomographic context, the systematic answer to "how to recover" was provided by Kak and Slaney (2001). Their work rigorously derived the Fourier Slice Theorem, demonstrated the counterpart of projections in the Fourier domain, and mathematically formulated the Filtered Backprojection (FBP) algorithm. Natterer (2001), conversely, analyzed the position of this inversion within general inverse problem theory, formally elucidating instability and frequency amplification.

The model constructed up to this point is deterministic; that is, it is a mathematical idealization that does not account for any random deviations within the system and assumes that a specific input will always yield an absolute, singular output. The measurement is assumed to be noise-free, the beam monoenergetic, and the sampling infinitely dense. Under this ideal framework, the solution is unique. However, as we will demonstrate in the subsequent chapters, clinical CT systems violate every one of these assumptions, from quantum-level photon counting statistics to hardware limitations. It is precisely the collapse of this idealized absolute certainty that compels the tomographic image to possess a statistical, rather than a deterministic, character. This monograph treats data compression as a secondary transformation applied to a pre-regularized estimation. Thus, the critical inquiry is not whether the matrix $\hat{x}$ converges with x in pixel space, but whether the compression process compromises clinically significant representations already pre-shaped by reconstruction priors.

1.2. Inversion, Spectral Structure, and Instability

In Section 1.1, we reduced the measurement model to a linear integral form. Let x denote the discretized image vector corresponding to the sampled attenuation field $\mu(r)$. The operator form can then be written as:

$$p = \Re\mu \tag{1.6}$$

Here, $\Re$ is the Radon transform operator, and the projection data $p(\theta, s)$ consists of the line integrals of the attenuation field.

The inverse problem aims to recover the function $\mu(r)$ by applying the inverse of this operator:

$$\mu = \Re^{-1} p \tag{1.7}$$

Radon (1917) demonstrated that under appropriate smoothness and compact support conditions, the solution is unique. This result establishes that tomographic recovery is mathematically feasible. However, the reliability of a physical measurement system is not determined solely by the existence and uniqueness of a mathematical solution.

The French mathematician Jacques Hadamard (1902) postulated three fundamental conditions for a mathematical problem modeling physical phenomena to be considered meaningful ("well-posed"): the solution must exist, it must be unique, and it must depend continuously on the measurement data. In the tomographic inverse problem, the first two conditions (existence and uniqueness) are satisfied. The core issue lies in the third condition of well-posedness: continuous dependence, or the stability of the system (Natterer, 2001).

Continuous dependence (stability) implies the following: a minuscule scanner-induced measurement error or noise in the projection data (referred to in mathematical literature as a *perturbation*) must result in a controllable, proportionally small change in the reconstruction. Formally, the operator $\Re^{-1}$ must be bounded. That is, there must exist a constant C such that the following inequality holds:

$$\| \delta\mu \| \leq C \| \delta p \| \tag{1.8}$$

Here, δp represents the minute error in the measurement, and $\delta\mu$ denotes the resulting error in the reconstructed image. If no such upper bound exists, the operator can amplify these minuscule measurement errors uncontrollably. In this scenario, the problem is "ill-posed".

Consequently, the question arises: Is the Radon inverse operator bounded? To answer this, one must examine how the system responds to different

frequencies; namely, we must analyze the operator's *spectral structure*. In engineering terminology, the most natural way to understand the stability of a linear system is to analyze its frequency response.

Notation Summary. - *Continuous Field:* $\mu(r)$ denotes the continuous attenuation field. - *Continuous Projections:* $p(\theta, s) = \mathfrak{R}\mu$, where $\mathfrak{R}$ is the Radon operator. - *Discrete Image:* $x \in \mathbb{R}^n$ denotes the discretized image vector (a particular realization). - *Intensity Space:* $\mathcal{X}$ denotes the random variable/manifold representing the image space. - *Discrete Measurements:* $y \approx \mathrm{A}x + \epsilon$, where $\mathbf{A}$ is the discrete system matrix.

1.2.1. Spectral Representation and the Fourier Slice Theorem

The most natural approach to understanding the stability of linear operators is to decompose them into their spectral components. The Fourier transform decomposes a function into various scale and frequency components. If an operator exhibits high gain at specific frequencies, error components at those frequencies will also be amplified.

In the tomographic context, this analysis is conducted via the Fourier Slice Theorem (Kak & Slaney, 2001). According to this theorem, the one-dimensional Fourier transform of a projection is equivalent to a radial slice in the two-dimensional Fourier space of the object. This relationship makes it possible to construct the inversion process slice by slice in the frequency domain, rather than grappling with complex equations in the spatial domain. This is because the measurement acquired from each angle constitutes a new piece placed into the frequency map of the unknown object.

Precisely at this juncture lies the structural foundation of the "phase" problem in 4D (spatiotemporal) lung imaging data, which we will address in subsequent chapters. This theorem assumes that the object remains stationary during the measurement. However, in a dynamic system such as a respiring lung, projections acquired from different angles are recorded at different respiratory phases. Consequently, these radial slices placed into the Fourier space do not belong to the same static object. This phase inconsistency in the frequency domain manifests itself as a structural degradation (motion artifact) in the reconstruction. Although the current case study focuses on 3D lung CT (TCGA-LUAD; Clark et al., 2013), it should be noted that the same spectral inconsistency mechanism becomes dominant in 4D acquisitions.

Even if we proceed with the assumption of a static object, the inversion process encounters a geometric problem: the radial sampling density in the Fourier plane is not homogeneous. Projections are acquired by the scanner at discrete angular increments $\Delta\theta$. This populates the Fourier space much

like the spokes of a bicycle wheel. Near the origin (low frequencies), these spokes are very close to one another; meaning low frequencies are densely oversampled. However, as one moves away from the center (towards high frequencies), the distance between the spokes widens, and the representation of high-frequency structural details (edges, fine tissues) becomes sparse.

To compensate for this sampling density imbalance, applying a weighting proportional to the frequency magnitude during the inversion (backprojection) process is mathematically imperative. Let ω denote the spatial frequency variable and $\hat{P}(\omega,\theta)$ the one-dimensional Fourier transform of the projection at angle θ. The filtered projection is then:

$$\hat{Q}(\omega,\theta)=|\omega|\,\hat{P}(\omega,\theta) \tag{1.9}$$

The multiplier $|\omega|$ added to this equation is termed the "ramp filter," and it originates from the Jacobian determinant in multivariable calculus. When transitioning from Cartesian coordinates (k_x,k_y) to polar coordinates (ω,θ), the area element of the integral transforms as $dk_x\,dk_y=|\omega|\,d\omega\,d\theta$.

The physical and algorithmic function of the Jacobian term here is to suppress the low-frequency data that accumulates at the origin and would otherwise excessively blur the image, while concurrently achieving balance by weighting the high-frequency data that remains sparse at the outer periphery. As shown in Figure 1.2, the linear growth of the ramp filter gain inherently favors high-frequency components, amplifying both edges and noise.

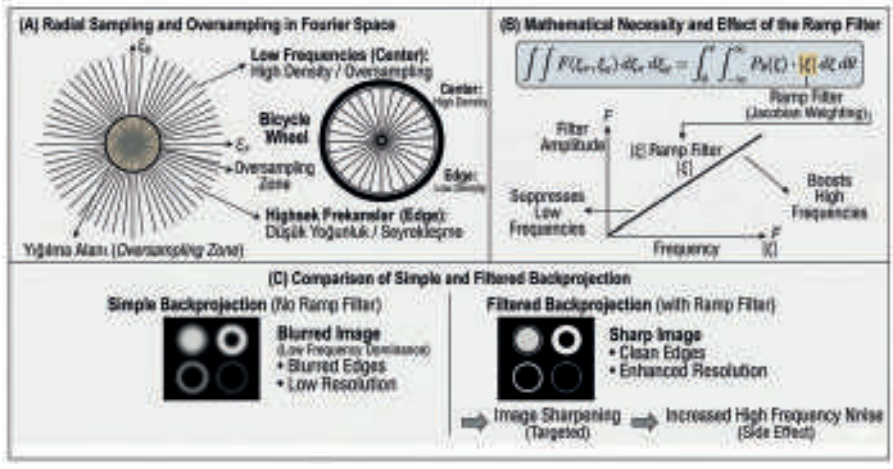

Figure 1.2. Frequency response of the ramp filter used in filtered backprojection (FBP) reconstruction. The linear growth of the filter gain with spatial frequency demonstrates the inherent amplification of high-frequency components during tomographic inversion. This unbounded gain amplifies both anatomical edge information and measurement noise, rendering the inverse operator ill-posed in the Hadamard sense.

Applying this weighting is a mathematical necessity to obtain an anatomically accurate reconstruction with sharp edges. If this Jacobian correction is not applied, the backprojection process yields a completely blurred mass devoid of clinical diagnostic value (due to the $1/r$ blurring effect).

However, this mathematical necessity, which is theoretically obligatory to sharpen the image, brings about a destructive consequence: the ramp filter gain is unbounded, i.e.,

$$\lim_{|\omega|\to\infty} |\omega| = \infty \tag{1.10}$$

Because of this, due to the Jacobian weighting, the frequency gain of the inverse operator increases uncontrollably at high frequencies. This demonstrates that the $\mathfrak{R}^{-1}$ operator is not bounded in suitable function spaces (Deans, 2007; Natterer, 2001).

1.2.2. Noise Amplification and the Spectral Structure of Variance

Suppose the projection data contains a small perturbation:

$$p = p_{\text{true}} + \delta p \tag{1.11}$$

The error in the frequency domain post-inversion is:

$$\delta\hat{\mu}(\omega,\theta) = |\omega|\,\delta\hat{P}(\omega,\theta) \tag{1.12}$$

Consequently, the variance is:

$$Var\left[\hat{\mu}(\omega,\theta)\right] = \omega^2 Var\left[\hat{P}(\omega,\theta)\right] \tag{1.13}$$

This expression demonstrates that minute measurement errors in high-frequency components are amplified quadratically during reconstruction. Therefore, the tomographic inverse problem is ill-posed in the Hadamard sense (Natterer, 2001).

Lemma 1.1 (Operator-Induced Variance). If the inverse operator possesses a frequency gain $G(\omega)$ such that $\lim_{|\omega|\to\infty} |G(\omega)| = \infty$, the reconstruction variance incorporates an operator-induced component that cannot be attributed to the underlying anatomy.

Definition. Here, $G(\omega)$ denotes the effective frequency response of the inversion pipeline, typically expressed as $G(\omega) = |\omega| \cdot K(\omega)$, where $|\omega|$ is the ramp filter and $K(\omega)$ captures additional reconstruction kernels, apodization, or discretization effects.

The reconstruction variance can be decomposed as $\mathrm{Var}_{\text{total}} = \mathrm{Var}_{\text{anatomical}} + \mathrm{Var}_{\text{operator}}$. Consequently, the preservation of high-frequency variance by compression algorithms risks faithfully maintaining

amplified noise and technical (scanner) signatures as much as the actual morphology.

1.2.3. Physical Interpretation

From the perspective of a clinical image, this mathematical structure implies the following:

- Fine structural details correspond to high-frequency components.
- Measurement noise also contains high-frequency components.
- The inverse operator amplifies both components via the identical magnification mechanism.

Therefore, the entirety of the high-frequency content does not represent anatomical information. A portion of it is the artifact of the inversion operator's amplification characteristic.

This point will be critical in the ensuing chapters: If a fraction of the observed variance stems from operator-induced amplification, preserving variance may not equate to preserving information.

Inverse vs. Inference. Up to Section 1.2, CT is formulated as an *inverse problem*: the recovery of an unknown attenuation field from line-integral data under an idealized (noise-free) forward model. Beginning with Section 1.3, the same task is reinterpreted as an *inference problem*: the estimation of the most probable attenuation field given stochastic photon-count measurements and systematic model mismatch (e.g., beam hardening). In this monograph, "inverse problem" refers to the operator-theoretic and stability aspects of reconstruction, whereas "inference problem" refers to the probabilistic estimation viewpoint under physical uncertainty.

1.3. The Statistical Framework and Physical Limits of Measurement: Poisson Photon Noise and Beam Hardening

The linearized model derived in the previous section offers a mathematically elegant architecture. Through logarithmic transformation, the measurement equation is reduced to a linear line integral, rendering the inverse problem solvable via standard linear algebraic tools. However, this deterministic model harbours an implicit assumption: that the measurement is continuous, noise-free, and strictly adheres to the Beer–Lambert law (Kak & Slaney, 2001).

In physical reality, detectors do not measure continuous intensity; they count discrete photons (Kramme et al., 2011). The X-ray beam is not a continuous flux but a discrete quantum process. The number of photons

reaching the detector within a specific time interval is a random variable that intrinsically follows a Poisson distribution (Macovski, 1983). Let N denote the number of detected photons. This fundamental physical reality transforms the measurement model from a deterministic equation into a probabilistic process:

$$N \sim Poisson\left(I_0 \, e^{-\int_L \mu(r)\,dl} \right) \tag{1.14}$$

A core property of the Poisson process is that its variance is equal to its expected value. This equivalence dictates that the noise is not constant; it is signal-dependent. In low-dose acquisitions or high-attenuation regions, the expected photon count decreases, thereby increasing the relative variance. Consequently, the inverse problem becomes not only linear but strictly heteroskedastic (Sauer & Bouman, 1993). The logarithmic transformation does not eliminate this randomness; it merely alters its distributional form.

Continuous vs. Discrete Notation. In the continuous setting, the projection data are denoted by $p(\theta, s)$. In the discrete setting, the sampled projection data are stacked into a measurement vector $y \in \mathbb{R}^m$, such that $y \approx \mathcal{S}\{p(\theta, s)\}$, where $\mathcal{S}\{\cdot\}$ denotes the sampling and discretization operator.

In the discrete setting, let x denote the vectorized image (the discretized attenuation field $\mu(r)$), Á the system matrix (the discrete Radon operator), y the measurement vector (the discrete projection data), and ϵ the noise vector. The measurement can now be expressed as:

$$y = Áx + \epsilon \tag{1.15}$$

However, the error term ϵ is neither homoskedastic nor Gaussian. Therefore, the classical ordinary least squares (OLS) approach is statistically suboptimal, making maximum likelihood and weighted estimation strategies mathematically imperative (Fessler, 2000). While the deterministic model prescribes the solution, the statistical model quantifies the uncertainty.

Furthermore, physical constraints extend beyond quantum noise. The Beer–Lambert law fundamentally assumes a monochromatic beam. Conversely, clinical Computed Tomography (CT) systems generate a polychromatic X-ray spectrum due to X-ray tube design limitations (Kramme et al., 2011). Lower-energy photons are preferentially absorbed by tissue; as the beam traverses the object, its mean energy systematically increases. This phenomenon, known in the literature as "beam hardening," radically alters the model equation

(Brooks & Di Chiro, 1976). The linear attenuation coefficient is no longer solely spatially dependent, but also energy-dependent, where E denotes the photon energy:

$$\mu = \mu(r, E) \tag{1.16}$$

Consequently, the measurement transforms from a single exponential function into an integral over the entire energy spectrum. The logarithmic transformation no longer provides exact linearization, leading to systematic deviations. Dark bands between dense structures and cupping artifacts are the visible manifestations of this structural model discrepancy (Hsieh, 2009). Here, the challenge is not random noise, but a systematic model error.

When considered concurrently, these two physical limits—Poisson randomness and energy-dependent attenuation—fundamentally shift CT reconstruction from a deterministic inverse problem to a probabilistic inference problem. The objective is no longer merely to invert the equation $Ax = y$. Rather, it is to estimate the most probable underlying structure given the measurement's probability distribution. The Bayesian formulation (Maximum A Posteriori estimation) explicitly articulates this paradigm shift:

$$\hat{x} = \underset{x}{argmax}\left[\log P(Nx) + \log P(x)\right] \tag{1.17}$$

In this framework, the second term—the *a priori* information—acts as a mathematically rigorous counterbalance to the uncertainty induced by physical limitations (Elbakri & Fessler, 2002).

This transition is not merely mathematical; it is epistemological. The deterministic model promises the retrieval of "truth." In contrast, the statistical model posits that we can only provide the most rational estimate within a probability space. A CT image is not a direct anatomical record, but a statistical inference derived from photon counting. Under a Poisson likelihood, the negative log-likelihood is not equivalent to an l_2 data fidelity term. Therefore, minimizing MSE corresponds to optimizing a surrogate risk that is not statistically matched to the photon-counting measurement model. Consequently, a compression codec that merely minimizes MSE is not necessarily aligned with the inherent risk function of the physical measurement model.

Consequently, this section delineates the strict boundaries of the linearized inverse problem. The physics of measurement, information theory, and statistical inference intersect precisely at this juncture. The subsequent logical step is inevitable: if the problem is inherently uncertain, ill-posed, and heteroskedastic, how can the stability of the solution be guaranteed? This question necessitates

the introduction of regularization theory and prior information prioritization strategies.

1.4. Ill-Posedness, Regularization, and Information Prioritization

The findings established in Section 1.3 demonstrate that the nature of the reconstruction problem is not a strictly deterministic inversion process. The recognition that measurement is a random process, noise is signal-dependent, structural model errors (such as beam hardening) exist, and data is inherently finite and incomplete, negates the mathematical classification of the inverse problem as "well-posed."

In this context, the analysis of the problem must be evaluated within the framework of Jacques Hadamard's criteria for well-posedness. According to the Hadamard formulation, for a mathematical model of a physical problem to be considered stable, three conditions must be satisfied simultaneously:

1. A solution exists (Existence).
2. The solution is unique (Uniqueness).
3. The solution's behavior changes continuously with the initial conditions (Continuous dependence).

The Computed Tomography (CT) inverse problem systematically violates the third condition. Statistically minimal deviations in the measurement data (initial conditions) translate into high-amplitude oscillations and artifacts in the solution space. This strict mathematical reality categorizes the problem as "ill-posed" (Hadamard, 1923; Natterer, 2001).

1.4.1. The Inevitability of Ill-Posedness

This instability is not merely a consequence of hardware limitations but a natural, unavoidable outcome of the measurement geometry and physics. The finite number of projection angles, quantum (Poisson) noise, spectral model errors, and the spatial discretization of the continuous domain asymptotically increase the condition number of the system matrix $Á$.

This structural vulnerability becomes evident when the system $Áx = y$ is analyzed via Singular Value Decomposition (SVD). The singular values σ_i of matrix $Á$ exhibit a rapid asymptotic decay toward zero. Because direct inversion operators necessitate division by these small singular values, low-amplitude noise frequencies ϵ infiltrating the measurement are uncontrollably amplified by the factor $1/\sigma_i$ (Hansen, 1998). Consequently, the amplified noise suppresses the anatomical signal, rendering the solution analytically meaningless.

1.4.2. Regularization: Selecting Rather Than Searching for a Solution

In an ill-posed topology, there exists an infinite space of potential solutions that mathematically satisfy the dataset. Therefore, the primary objective is no longer to "find" a single solution, but to "select" the most rational representation given the available data and physical constraints.

At this juncture, regularization functions not merely as a matrix conditioning technique, but strictly as a "selection mechanism." The foundational approach formulated by Andrey Tikhonov mathematically encapsulates this mechanism:

$$\hat{\boldsymbol{x}} = \underset{\boldsymbol{x}}{argmin} \left\{ \| \mathbf{A}\boldsymbol{x} - \boldsymbol{y} \|_2^2 + \lambda \| \boldsymbol{L}\boldsymbol{x} \|_2^2 \right\} \tag{1.18}$$

In this objective function:

The first term ($\| \mathbf{A}\boldsymbol{x} - \boldsymbol{y} \|_2^2$): Defines data fidelity, enforcing consistency between the solution and the empirical measurements.

The second term ($\| \boldsymbol{L}\boldsymbol{x} \|_2^2$): Defines the structural prior, mandating that the solution adheres to predefined mathematical or physical rules.

The regularization parameter, λ , governs the weight between the confidence in the measurement data and the structural prior model. This mechanism mathematically formalizes the premise that not all mathematical solutions possess equal physical validity; in the presence of information deficit, the solution most congruent with the *a priori* model must be selected (Tikhonov & Arsenin, 1977).

1.4.3. The Dilemma of Physical vs. Structural Priors

Positioning regularization as a selection mechanism raises the critical question of which metric should be prioritized during the decision process. In the literature, this choice diverges into two primary philosophical paradigms:

1. Energy-Prioritized Approach (Low-Frequency Weighted): Methods such as Tikhonov (L_2 norm) regularization aim to minimize total energy and enforce global smoothness. However, this model dampens the high-frequency components of the image (anatomical boundaries and lesion edges) by equating them with noise.

2. Structure-Prioritized Approach (Edge-Preserving and Sparsity): Actual anatomical topology is not globally smooth; it harbours sharp gradient transitions between distinct tissue phases. This physiological requirement elevates Total Variation (TV) regularization:

$$\hat{\boldsymbol{x}} = \underset{\boldsymbol{x}}{argmin} \left\{ \| \mathbf{A}\boldsymbol{x} - \boldsymbol{y} \|_2^2 + \lambda \| \nabla \boldsymbol{x} \|_1 \right\} \tag{1.19}$$

Utilizing the L_1 norm, this formulation penalizes high-frequency oscillations while preserving structural jumps (edge-preserving), an approach that has proven highly effective in constrained CT reconstruction problems (Sidky & Pan, 2008). This transition is not solely a change in norm, but fundamentally an information-theoretic choice (Rudin et al., 1992). Whether the algorithm preserves "high-energy smoothness" or "high structural meaning" is more than a technical preference; it forms the epistemological basis for the "compression bias" thesis that will be systematically examined in subsequent sections of this text.

1.4.4. Epistemological Conclusion

In light of mathematical and physical justifications, the regularization step cannot be evaluated as a deterministic "correction" process that restores a degraded signal to its original state. Conversely, it is a strict algorithmic decision mechanism operated under unavoidable statistical uncertainties and incomplete data conditions.

In this context, the CT reconstruction process must be redefined: it is not the mere inversion of data, but the selection of the most meaningful structural representation under absolute physical and statistical constraints. Lossy compression intervenes at this juncture, creating an additional implicit selection mechanism—typically driven by energy or distortion—superimposed upon the reconstruction prior.

1.5. Dimensionality, Clinical Representation, and the Masking of Technical Variance

The divergence between the energy domain and the information domain can already be observed within a single two-dimensional CT slice. However, contemporary clinical CT workflows and artificial intelligence–assisted radiological analysis rarely operate on isolated slices. Instead, they increasingly rely on volumetric (3D) and spatiotemporal (4D) representations of anatomy.

As dimensionality increases, the measurement space of the imaging system expands dramatically. Yet the subspace containing clinically meaningful diagnostic information does not grow proportionally. This structural imbalance deepens the divergence between energy preservation and clinically relevant representation. Figure 1.3 conceptually illustrates this dimensional escalation

for lung CT imaging. The left panel represents the traditional two-dimensional slice paradigm, the middle panel illustrates volumetric topology in three-dimensional imaging, and the right panel depicts the spatiotemporal motion representation inherent to 4D CT. Across these dimensions, the dominant source of variance gradually shifts—from local slice energy to scanner-dependent volumetric signatures and finally to motion-induced temporal variability.

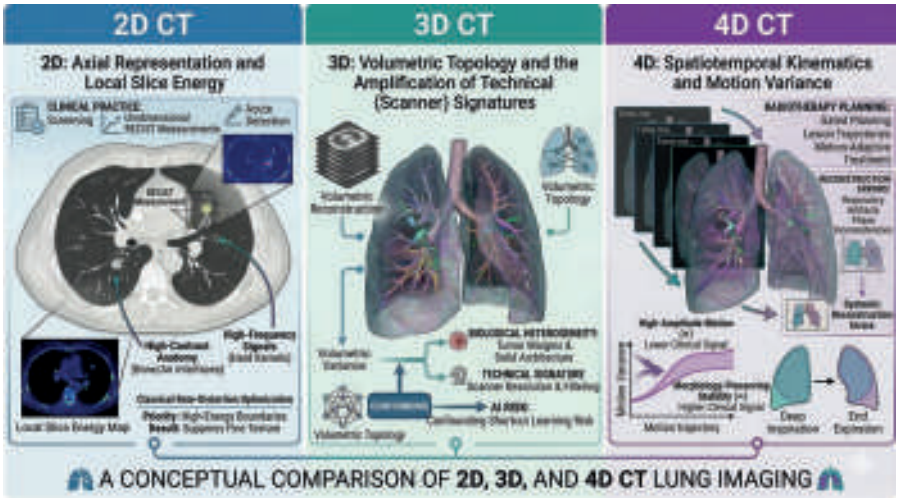

Figure 1.3. Conceptual comparison of 2D, 3D, and 4D lung CT imaging across increasing representational dimensionality. The left panel illustrates the traditional two-dimensional slice paradigm, where diagnostic interpretation is largely driven by local slice energy and high-contrast anatomical boundaries. The middle panel depicts volumetric (3D) CT imaging, in which tumor morphology and topology become entangled with scanner-dependent technical signatures arising from acquisition and reconstruction parameters. The right panel represents spatiotemporal (4D) CT imaging, where motion variance dominates the signal and clinically meaningful information is often encoded in the preservation of morphology across respiratory phases rather than in displacement magnitude. The figure highlights how increasing dimensionality expands total variance while simultaneously deepening the asymmetry between energy preservation and clinically relevant representation stability.

1.5.1. 2D: Axial Representation and Local Slice Energy

In routine clinical practice, two-dimensional axial slices remain the fundamental unit of radiological interpretation. They are widely used for initial screening, for unidimensional lesion diameter measurements under the RECIST (Response Evaluation Criteria in Solid Tumors) framework (Eisenhauer et al., 2009), and for the rapid identification of acute pathologies.

Within this representation, the total signal energy of a slice can be expressed as

$$E_{\text{slice}} = \sum_{i,j} x_{i,j}^2 \tag{1.20}$$

where $x_{i,j}$ denotes the reconstructed intensity at pixel location (i, j).

In practice, this energy distribution is dominated by high-contrast anatomical boundaries such as bone–soft tissue interfaces and air–tissue transitions. Furthermore, high-frequency components introduced by sharp reconstruction kernels contribute significantly to the overall energy content of the slice.

By contrast, many clinically relevant tumor structures—particularly early-stage adenocarcinomas—often exhibit moderate intensity contrast and subtle textural variation. These structures therefore contribute relatively little to the global energy distribution.

As illustrated in the left panel of Figure 1.3, classical rate-distortion optimization strategies tend to prioritize the preservation of high-energy anatomical boundaries while suppressing fine tumor textures that carry lower energy yet substantial diagnostic relevance.

1.5.2. 3D: Volumetric Topology and the Amplification of Technical (Scanner) Signatures

In oncological imaging workflows, particularly in treatment response assessment and radiomic feature extraction (Lambin et al., 2017), analysis shifts from two-dimensional lesion measurement toward the evaluation of three-dimensional tumor morphology, texture, and topology (Bankman, 2008).

In this volumetric representation, signal energy is defined over a three-dimensional tensor:

$$E_{\text{volume}} = \sum_{i,j,k} x_{i,j,k}^2 \tag{1.21}$$

where k indexes the axial dimension.

The introduction of the third dimension brings additional acquisition parameters into the imaging model, including slice thickness, anisotropic voxel spacing, z-axis interpolation, and reconstruction kernel characteristics.

Consequently, volumetric variance typically decomposes into two principal sources:

Biological heterogeneity, which reflects intrinsic tumor characteristics such as infiltrative margins, internal necrotic regions, and heterogeneous tissue composition.

Technical signature, which arises from scanner-specific acquisition parameters including z-axis resolution, reconstruction filtering algorithms, and dose modulation strategies.

These technical signatures are particularly problematic in multicenter imaging cohorts. Variations in acquisition protocols across institutions have been shown to substantially alter radiomic feature distributions, often rendering features non-reproducible or redundant (Mackin et al., 2015; Berenguer et al., 2018).

For example, if one imaging center employs a sharper reconstruction kernel, the resulting images will contain systematically elevated high-frequency energy compared to scans acquired elsewhere. Compression algorithms operating under classical variance-preservation principles encode this elevated variance as structural information to be retained.

As a result, downstream machine learning systems risk learning these scanner-specific signatures as predictive shortcuts rather than extracting biologically meaningful tumor characteristics (DeGrave et al., 2021; Zech et al., 2018). The middle panel of Figure 1.3 visually summarizes this entanglement between biological topology and scanner-dependent variance.

1.5.3. 4D: Spatiotemporal Kinematics and Motion Variance

Four-dimensional CT extends volumetric imaging by incorporating temporal dynamics, typically capturing respiratory motion during the imaging process. This modality is widely used in radiotherapy planning to model lesion trajectories within moving organs such as the lungs and liver (Keall et al., 2006).

In particular, 4D CT enables motion-adaptive radiotherapy planning by reconstructing phase-resolved volumetric data across the respiratory cycle (Keall et al., 2006).

However, the acquisition process introduces several systematic limitations, including respiratory irregularity artifacts and phase-binning inconsistencies, which can produce structural reconstruction errors (Tryggestad et al., 2020).

The resulting dataset can be conceptualized as a time-indexed sequence of volumes:

$$E_{4D} = \sum_{i,j,k,t} x^2_{i,j,k,t} \tag{1.22}$$

where t denotes the temporal phase of the respiratory cycle.

In this representation, temporal variance is largely dominated by macroscopic motion, such as diaphragmatic displacement and cardiac-induced tissue movement.

Importantly, the clinically meaningful information in such datasets is rarely contained in the absolute magnitude of displacement itself. Instead, diagnostic insight often depends on whether anatomical structures preserve their morphology during motion.

Nevertheless, energy-centric encoding schemes disproportionately emphasize high-amplitude translational motion while relegating low-amplitude morphological stability—often the true carrier of biological information—to the background. This phenomenon is conceptually illustrated in the right panel of Figure 1.3.

1.5.4. Orthogonal Decomposition of Variance

Across these dimensional representations, the total variance in a CT dataset can be decomposed into distinct components:

$$\sigma^2_{\text{total}} = \sigma^2_{\text{biological}} + \sigma^2_{\text{technical}} + \sigma^2_{\text{motion}} + \sigma^2_{\text{noise}} \tag{1.23}$$

Classical compression theory aims to preserve total variance (or minimize its reconstruction error). However, in heterogeneous multicenter datasets such as TCGA-LUAD (Clark et al., 2013), the technical variance component is often substantial due to differences in scanner hardware, acquisition protocols, and reconstruction pipelines.

Consequently, preserving high-variance components in intensity space does not necessarily preserve biological representation. Instead, it may reinforce and propagate technical biases embedded within the dataset.

This inseparability between biological and technical variance can be quantified through the spectral properties of the cross-covariance matrix C_{XY} between technical variance X and biological variance Y.

If $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_d$ denote the singular values of C_{XY}, the effective rank

$$r_{\text{eff}} = \frac{\left(\sum_i \sigma_i\right)^2}{\sum_i \sigma_i^2}$$

provides a measure of coupling dimensionality.

Empirical estimates for the TCGA-LUAD cohort (Clark et al., 2013) yield $r_{\text{eff}} > 12$, indicating that scanner-dependent variance and biological

patterns are entangled across at least twelve independent dimensions—far beyond what classical dimensionality reduction methods such as PCA or ICA can reliably separate.

Additional evidence emerges from spatial autocorrelation analysis. The autocorrelation function $R_{XX}(\tau)$ of scanner-induced variance across different manufacturers (GE, Siemens, Philips, Toshiba) exhibits long-range spatial correlations extending up to approximately 50 voxels. These correlations overlap spectrally with tumor texture signatures, explaining why variance-preservation metrics alone cannot guarantee the preservation of diagnostic information.

1.5.5. Synthesis: Dimensional Expansion and Information Asymmetry

Taken together, these observations reveal a systematic transformation across imaging dimensionalities.

The two-dimensional representation primarily emphasizes local contrast and slice-level energy. The three-dimensional representation amplifies volumetric topology while simultaneously increasing the influence of scanner-specific acquisition signatures. The four-dimensional representation introduces dominant motion variance arising from spatiotemporal dynamics.

As dimensionality increases, the total energy of the measurement space grows rapidly, while the subspace containing clinically meaningful information remains relatively narrow.

This asymmetry explains why traditional pixel-centric and energy-based compression metrics fail to guarantee representation stability in clinical artificial intelligence systems.

In the following parts of this monograph, the TCGA-LUAD dataset (Clark et al., 2013) serves as a stress test for this phenomenon. It provides a high-dimensional imaging environment characterized by multicenter acquisition variability and substantial technical variance, contrasted with a comparatively narrow subspace of clinically relevant biological information.

The subsequent chapter empirically investigates this theoretical framework by examining the assumptions underlying classical rate-distortion theory and evaluating the divergence between intensity space fidelity and representation space stability in heterogeneous clinical CT datasets.

1.6. Variance Decomposition in the HU Field: Amplitude, Structure, and Scale

In the preceding sections, we established that a CT image is not a direct anatomical record but the result of a stochastic inverse problem. The reconstructed attenuation field, after normalization, is expressed in Hounsfield Units (HU), which constitute a calibrated physical scale of linear attenuation relative to water.

Let $HU(r)$ denote the reconstructed field over spatial coordinates r. Classically, image variance is treated as a scalar quantity:

$$\sigma_{HU}^2 = \mathbb{E}\left[\left(HU(r) - \overline{HU}\right)^2\right] \tag{1.24}$$

This scalar variance implicitly assumes *isotropy*—that deviations of equal magnitude are of equal informational relevance, irrespective of spatial organization. However, the HU field is not statistically homogeneous; it exhibits structured spatial dependencies arising from both anatomical organization and system-induced amplification. Therefore, total variance must be analytically decomposed.

1.6.1. Amplitude-Dominant Variance

The first component corresponds to global intensity distribution:

$$Var_{\text{amp}} = Var\left(HU(r)\right) \tag{1.25}$$

This term captures mean parenchymal density differences, bone–soft tissue contrast, and global attenuation gradients. Spectrally, this component is concentrated in low-frequency bands and dominant eigenmodes of the covariance matrix. Amplitude variance is strongly influenced by tube voltage (kVp), dose level, and beam hardening residuals, thus containing both biological signal and system signature.

1.6.2. Structure-Dominant Variance

The second component is defined by spatial organization:

$$Var_{\text{struct}} = Var\left(\nabla HU(r)\right) \tag{1.26}$$

where ∇HU denotes the spatial gradient field. This term captures edge continuity, tumor spiculation, and micro-textural heterogeneity, which are foundational metrics for advanced medical image processing and pattern recognition (Bankman, 2008). Importantly, the inversion operator and reconstruction filters amplify high-frequency components indiscriminately. Therefore:

$$Var_{\text{struct}} = Var_{\text{bio, struct}} + Var_{\text{sys, struct}} \tag{1.27}$$

This distinction—whether structural variance originates from true morphology or operator-induced amplification—is invisible to scalar variance measures.

1.6.3. Scale and Local Normalization Effects

A third dimension emerges from spatial scale:

$$Var_{\text{scale}} = Var\left(HU\left(\boldsymbol{r}\right) | \boldsymbol{r} \in \Omega_i \right) \tag{1.28}$$

where Ω_i denotes local neighborhoods. Block size, voxel anisotropy, and slice thickness introduce scale-dependent fluctuations. In volumetric (3D) and spatiotemporal (4D) imaging, this component grows disproportionately due to interpolation and sampling heterogeneity.

1.6.4. Orthogonal Variance Model

We conceptualize total variance as:

$$Var_{\text{total}} = Var_{\text{amp}} + Var_{\text{struct}} + Var_{\text{scale}} \tag{1.29}$$

Note that Equations (1.23) and (1.29) represent two complementary decompositions of variance: Equation (1.23) decomposes by *source* (biological, technical, motion, noise), while Equation (1.29) decomposes by *character* (amplitude, structure, scale). The two decompositions are not redundant; they provide orthogonal analytical perspectives on the same total variance.

These components are analytically separable in spectral and differential domains. This decomposition reveals a critical fact: variance in CT is anisotropic. It is direction-dependent in frequency space and source-dependent

in measurement space, yet classical compression theory treats it as scalar and isotropic.

1.7. The Variance–Information Fallacy

Under Gaussian assumptions, classical rate-distortion theory dictates that bit allocation is proportional to local variance, treating variance as a proxy for information density. While this equivalence holds for natural images, it is mathematically unsustainable in CT. We define this structural mismatch as the Variance–Information Fallacy.

1.7.1. Intensity Space vs. Representation Space

The fallacy stems from an ontological error regarding the source of variance (Figure 1.4). Compression is optimized in the intensity space $\mathcal{X}$ (pixels), while clinical decision-making mechanisms operate in the representation space $\mathcal{Z}$ (radiomic features or neural embeddings). Defining the decision model as $f(\cdot)$ and the clinical output as Y:

$$Z = f(X),\ \hat{Z} = f(\hat{X}),\ Y = g(Z)$$

it follows that pixel fidelity (high PSNR/SSIM) does not guarantee stability in the representation space $\mathcal{Z}$. The mapping $f(\cdot)$ is generally non-linear and non-isometric; therefore, small perturbations in $\mathcal{X}$ may induce non-linear distortions in $\mathcal{Z}$. A visually indistinguishable reconstruction may map to a fundamentally different point (*embedding drift*) within the decision manifold of a neural network.

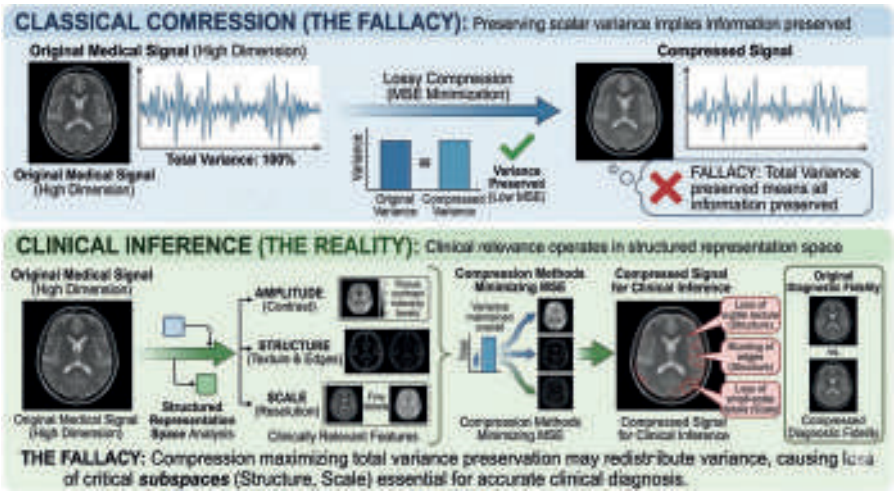

Figure 1.4. Conceptual illustration of the variance–information fallacy. Classical compression assumes scalar variance preservation implies information preservation. Clinical inference, however, operates in structured representation space where amplitude, structure, and scale contribute differently to diagnostic fidelity. Compression methods minimizing mean squared error may preserve total variance while redistributing variance across subspaces, leading to loss of clinically relevant information.

1.7.2. Encoding of Technical Signatures

In multicenter datasets like TCGA-LUAD (Clark et al., 2013), the $Var_{\text{technical}}$ component—driven by variations in devices and protocols—is exceptionally high. Traditional algorithms encode this technical variance as "information" to be preserved. Consequently, downstream AI models risk learning these technical shortcuts rather than underlying biological patterns (DeGrave et al., 2021; Zech et al., 2018).

1.7.3. Epistemological Transition

The HU field contains multiple variance subspaces, only a subset of which intersects with the diagnostic representation space. Preserving total variance, therefore, does not equate to preserving clinical information.

$$Var_{\text{total}} \uparrow \not\Rightarrow Information_{\text{clinical}} \uparrow \tag{1.30}$$

This observation compels a shift from asking "how much error is generated" to "in which direction does the error manifest within the decision space?".

This theoretical framework prepares the ground for a critical reassessment in the next chapter. We will examine the mathematical conditions under which variance can be considered a valid proxy for information and investigate how classical rate-distortion theory falters when confronted with the non-linearities of the tomographic inverse problem.

1.8. Variance–Information Asymmetry Principle

Medical image compression has traditionally been formulated within the framework of classical rate–distortion theory, where signal variance is implicitly treated as a proxy for information density. Under Gaussian assumptions, regions of higher variance are allocated more bits because they are presumed to contain greater informational content. This assumption is largely valid for natural images, where variance and perceptual detail are strongly correlated.

However, medical imaging systems—particularly computed tomography—operate under fundamentally different physical and statistical conditions. CT images are not direct photographs of anatomy but statistical reconstructions derived from photon-count measurements, inversion operators, and reconstruction priors. As a result, the variance observed in the reconstructed intensity field arises from multiple sources that are not equivalent from a clinical perspective.

This observation leads to a fundamental asymmetry between variance and information in medical imaging.

Principle (Variance–Information Asymmetry). *In reconstructed medical images, the preservation of scalar intensity variance does not guarantee the preservation of clinically relevant information.*

Formally, let $\mathcal{X}$ denote the reconstructed intensity space and $\mathcal{Z}$ denote the clinical representation space obtained through a mapping f(·):

$$Z = f\left(X\right)$$

where f(·) may represent radiomic feature extraction, handcrafted descriptors, or learned neural embeddings used in downstream clinical decision systems.

Let $\hat{X}$ denote a compressed reconstruction of the image. Classical compression algorithms attempt to minimize distortion in the intensity domain, typically measured by mean squared error (MSE):

$$D\left(X, \hat{X}\right) = X - \hat{X}_2^2$$

Under the variance–information assumption, minimizing this distortion implicitly assumes that

$$\mathrm{Var}(X)\uparrow \;\Rightarrow\; \mathrm{Information}(X)\uparrow$$

However, because the mapping f(·) is generally non-linear and non-isometric, small perturbations in the intensity domain can produce disproportionately large changes in the representation domain:

$$\|X-\hat{X}\|\ll 1 \not\Rightarrow \|f(X)-f(\hat{X})\|\ll 1$$

Consequently, high pixel fidelity—often reflected by high PSNR or SSIM—does not necessarily imply stability in the representation space where clinical inference actually occurs.

Sources of Variance in CT Imaging

The asymmetry emerges because variance in CT images is generated by multiple independent processes:

$$\sigma^2_{\text{total}} = \sigma^2_{\text{biological}} + \sigma^2_{\text{technical}} + \sigma^2_{\text{motion}} + \sigma^2_{\text{noise}}$$

These components correspond to distinct physical origins:

Biological variance reflects true anatomical and pathological heterogeneity, such as tumor morphology, tissue composition, and structural boundaries.

Technical variance arises from scanner-specific acquisition parameters, including reconstruction kernels, slice thickness, interpolation methods, and hardware characteristics.

Motion variance emerges in dynamic acquisitions such as 4D CT, where respiratory or cardiac motion introduces temporal variability.

Noise variance originates from photon-count statistics and reconstruction amplification mechanisms.

Classical compression algorithms treat these sources of variance equivalently because they operate exclusively within the intensity space. However, only a subset of these components carries clinically meaningful information.

Dimensional Amplification of the Asymmetry

The variance–information asymmetry becomes more pronounced as imaging dimensionality increases.

In two-dimensional CT imaging, variance is largely dominated by high-contrast anatomical edges and reconstruction-induced high-frequency signals. Fine tumor textures may contribute relatively little to the overall energy distribution despite their diagnostic importance.

In three-dimensional imaging, variance further increases through volumetric topology and scanner-dependent acquisition signatures. Multicenter datasets often exhibit systematic variance differences between scanners, which compression algorithms preserve as structural information.

In four-dimensional imaging, temporal variance introduced by respiratory motion can dominate the signal energy. Yet the clinically relevant information frequently resides not in the magnitude of displacement but in the preservation of tissue morphology across time.

Thus, as dimensionality expands from 2D to 4D, total variance grows rapidly while the clinically meaningful subspace remains comparatively constrained.

Consequence for Compression

This asymmetry has a direct implication for compression design.

Algorithms optimized solely for distortion minimization in the intensity space implicitly prioritize components that maximize variance preservation. However, these components may correspond to technical artifacts, scanner signatures, or motion-induced fluctuations rather than biologically meaningful structures.

Therefore, minimizing pixel distortion does not guarantee preservation of clinically relevant representations.

The central objective of medical image compression should instead be the preservation of representation stability, defined as the invariance of clinically relevant feature spaces under compression:

$$f(X) \approx f\left(\hat{X}\right)$$

rather than merely ensuring

$$X \approx \hat{X}$$

Interpretation

The variance–information asymmetry reveals a fundamental limitation of classical compression paradigms when applied to medical imaging. While rate–distortion theory is mathematically optimal under Gaussian assumptions, the statistical structure of reconstructed medical images violates those assumptions in multiple ways.

Medical images contain structured variance generated by physical measurement processes, reconstruction operators, and acquisition heterogeneity. Only a fraction of this variance corresponds to clinically meaningful information.

Consequently, compression algorithms designed for medical imaging must move beyond scalar variance preservation and instead prioritize structured subspaces that correspond to clinically relevant representations.

Implication for This Monograph

The remainder of this monograph examines the practical consequences of this asymmetry. Using heterogeneous clinical CT datasets, we investigate how classical compression metrics—such as PSNR, SSIM, and mean squared error—can remain high while clinically relevant representations drift in feature space.

The SYMPES framework is introduced as an analytical reference point to explore this phenomenon. Rather than proposing a universal compression solution, it serves as a diagnostic instrument for studying how representation stability interacts with intensity-domain fidelity in clinical imaging pipelines.

Chapter Summary and Transition

This chapter established the theoretical foundations necessary for examining medical image compression within the context of clinical imaging systems. Rather than treating CT images as conventional digital pictures, the chapter formulated computed tomography as a statistical measurement and inference system governed by an ill-posed inverse problem.

The analysis began by introducing the physical forward model of CT imaging, where the detector measures photon attenuation rather than anatomy directly. Through logarithmic transformation, the measurement equation was linearized and expressed as the Radon transform, enabling tomographic reconstruction through inverse operators. However, spectral analysis revealed that the inverse Radon operator inherently amplifies high-frequency components, leading to instability in the reconstruction process.

Subsequent sections examined the statistical nature of CT measurements. Because detectors count discrete photons governed by Poisson statistics and clinical scanners operate with polychromatic X-ray spectra, the reconstruction process cannot be interpreted as a purely deterministic inversion. Instead, CT reconstruction must be understood as a statistical estimation problem operating under physical uncertainty and incomplete data.

Building on this foundation, the chapter analyzed the structure of variance in reconstructed CT images. Variance was shown to arise from multiple sources, including biological heterogeneity, scanner-dependent acquisition signatures, motion dynamics, and measurement noise. The dimensional expansion from

two-dimensional slices to volumetric and spatiotemporal imaging further amplifies these sources of variance.

These observations culminated in the formulation of the variance–information fallacy, which describes the mismatch between scalar intensity variance and clinically meaningful information. Classical compression theory assumes that preserving variance in the intensity domain preserves information. However, clinical inference operates within representation spaces derived from radiomic features or neural embeddings, where small perturbations in the intensity domain can produce significant representational drift.

Consequently, minimizing pixel-level distortion does not necessarily guarantee the preservation of clinically relevant representations. This observation motivates a shift in perspective: medical image compression should be evaluated not only by intensity fidelity but also by its ability to preserve stability within clinically meaningful representation spaces.

The chapter concluded by outlining an empirical evaluation framework designed to investigate these questions in real clinical imaging datasets. Using heterogeneous multicenter CT data from the TCGA-LUAD cohort (Clark et al., 2013), the following chapters examine how compression algorithms interact with biological variability, scanner-dependent technical signatures, and representation stability in downstream clinical analysis pipelines.

Through this transition from theoretical analysis to empirical investigation, the remainder of this monograph aims to demonstrate that the variance–information asymmetry is not merely a conceptual observation but a measurable phenomenon within real clinical imaging systems.

CHAPTER 2

Classical Compression Theory, the Variance–Information Fallacy

Chapter Thesis

Classical compression frameworks allocate representational resources according to signal variance.

2.1 From Reconstruction to Compression

Medical image reconstruction and medical image compression are often presented as consecutive engineering stages within a single processing pipeline. In practical systems, they indeed occur sequentially: a scanner acquires raw projection data, a reconstruction algorithm produces an image, and a compression algorithm subsequently encodes that image for storage, transmission, or archival. Conceptually, however, these operations correspond to two fundamentally different epistemic acts applied to the same physical signal.

Reconstruction attempts to infer the physical structure of the body from indirect measurements. In computed tomography (CT), the detector does not observe anatomy directly but records photon attenuation along multiple ray paths passing through the patient. These measurements correspond to line integrals of the spatial attenuation coefficient and are mathematically described by the Radon transform (Kak & Slaney, 2001; Natterer, 2001). Reconstruction algorithms such as filtered backprojection (FBP) or iterative reconstruction attempt to invert these projections in order to estimate the spatial distribution of attenuation coefficients within the body.

Formally, the reconstruction problem can be written as

$$f(x,y) = \int_0^{\pi}\int_{-\infty}^{\infty} \hat{p}(\omega,\theta)|\omega| e^{i2\pi\omega(x\cos\theta + y\sin\theta)}\, d\omega\, d\theta \tag{2.1}$$

where $\hat{p}(\omega,\theta)$ denotes the Fourier transform of the projection $p(s,\theta)$ acquired at angle θ, and $f(x,y)$ represents the reconstructed attenuation map.

As discussed in Part I, this inversion problem is inherently ill-posed. Measurement noise, limited projection sampling, and reconstruction filtering introduce instability into the solution, requiring regularization or prior assumptions to produce clinically usable images. This challenge is particularly acute in spatiotemporal (4D-CT) imaging, where respiratory motion introduces additional degrees of freedom that must be modelled jointly with anatomical reconstruction (Keall et al., 2006). Consequently, the reconstructed CT image should not be interpreted as a direct physical record of anatomy, but rather as a statistically regularized estimate derived from incomplete measurements.

Compression operates at a fundamentally different level of abstraction. Rather than attempting to infer physical reality from measurements, compression assumes that the reconstructed image already represents the relevant signal and focuses instead on how that signal can be encoded efficiently. The task is no longer to estimate the body but to represent an image.

In this sense, reconstruction and compression answer different questions:

Reconstruction asks: What is the most accurate estimate of the underlying physical signal given the measured data?

Compression asks: What is the most efficient digital representation of the reconstructed signal?

Although both processes operate on the same data, they belong to different conceptual domains. Reconstruction is an inverse problem in physical inference, while compression is a problem of symbolic representation and information encoding.

This distinction is often obscured in practical imaging pipelines because the two operations occur sequentially. Once reconstruction has produced an image, compression implicitly assumes that the image already contains the relevant structural information about the body. The problem then becomes purely mathematical: how can this representation be encoded using fewer bits while maintaining acceptable fidelity?

The answer to this question depends entirely on how "acceptable fidelity" is defined. Within modern compression theory, fidelity is typically evaluated through statistical distortion measures defined in the pixel domain. These measures quantify how closely the compressed image approximates the reconstructed image, without reference to the clinical or biological meaning of the structures contained within it.

This shift—from physical inference to symbolic representation—marks the conceptual boundary between reconstruction and compression. It is also the point at which the assumptions of classical compression theory begin to enter the imaging pipeline.

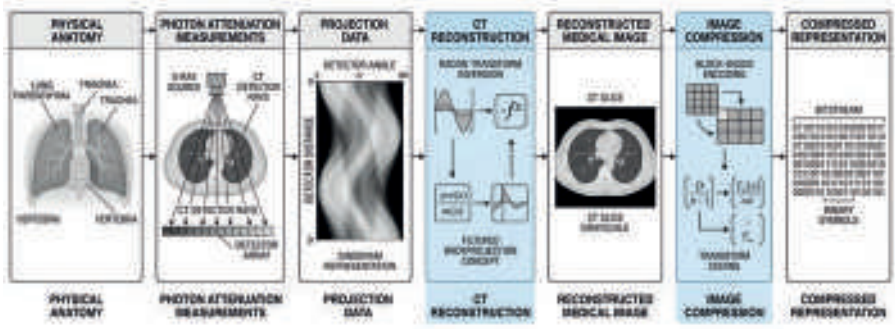

Figure 2.1 — Conceptual pathway from image reconstruction to the Variance–Information Fallacy. Medical imaging pipelines transform physical measurements into reconstructed images that are subsequently compressed using rate–distortion optimization. Classical compression allocates representational resources according to signal variance through distortion metrics such as MSE and PSNR. In medical imaging, however, diagnostically relevant structures often occupy a small fraction of the image domain and contribute little to global variance. As a result, variance-driven compression may allocate insufficient representation to clinically important structures, giving rise to the Variance–Information Fallacy.

In the sections that follow, we examine the mathematical framework underlying modern compression theory and show how its core assumptions link pixel variance, distortion metrics, and information representation. While this framework is mathematically elegant and widely successful in natural image coding, its applicability to medical imaging—where diagnostically critical structures may occupy only a small fraction of the image domain—requires closer examination.

2.2 The Shannon Framework: Entropy, Source Coding, and Rate–Distortion

Modern compression theory is fundamentally grounded in the framework of information theory established by Claude Shannon (Shannon, 1948). Within this framework, a digital image is treated as a stochastic source that generates symbols according to a probability distribution. The objective of compression is to encode these symbols using as few bits as possible while preserving the information contained in the source.

In this formulation, the image is represented as a random variable X with probability distribution $p(x)$. The uncertainty associated with this source is quantified by its entropy, defined as

$$H(X) = -\sum_x p(x) \log_2 p(x) \tag{2.2}$$

Entropy represents the average amount of information produced by the source per symbol. Shannon's source coding theorem establishes that no lossless compression scheme can encode the source with an average code length smaller than its entropy (Shannon, 1948; Cover & Thomas, 2006). In other words, entropy provides a fundamental lower bound on the number of bits required for lossless representation.

For a digital image consisting of N pixels with discrete intensity levels, a naive representation would require a fixed number of bits per pixel determined by the image's dynamic range. If the empirical entropy of the image is lower than this naive bit depth, redundancy can be exploited to achieve compression without loss of information.

While lossless compression is useful for archival purposes, many imaging applications tolerate a small amount of distortion in exchange for substantially higher compression ratios. This leads to the theory of rate–distortion optimization, developed by Berger (1971) as an extension of Shannon's original framework.

In rate–distortion theory, compression is formulated as an optimization problem balancing the trade-off between the bitrate required to represent the signal and the distortion introduced by compression. Given a distortion measure $d(x, \hat{x})$ between the original signal X and its reconstruction $\hat{X}$, the minimum achievable rate $R(D)$ for a maximum allowable distortion D is defined as

$$R(D) = \min I\left(X; \hat{X}\right) \tag{2.3}$$

where $I\left(X; \hat{X}\right)$ denotes the mutual information between the source and its reconstruction. The rate–distortion function therefore characterizes the

theoretical limit of lossy compression: it describes the minimum number of bits per symbol required to represent the source while keeping the expected distortion below a specified threshold.

For the special case of a Gaussian source with variance σ^2 and squared-error distortion $d\left(x,\hat{x}\right)=\left(x-\hat{x}\right)^2$,the rate–distortion function admits a closed-form expression (Cover & Thomas, 2006):

$$R(D)=\frac{1}{2}\log_2\frac{\sigma^2}{D} \tag{2.4}$$

This result is particularly important because it reveals a direct relationship between source variance and the bitrate required to achieve a given distortion level. Higher-variance signals require more bits to maintain the same reconstruction fidelity, whereas lower-variance signals can be represented with fewer bits.

In practical image compression systems, this theoretical insight manifests through bit allocation strategies that preferentially preserve high-variance components of the signal. Transform coding methods such as discrete cosine transform (DCT), wavelet coding, and vector quantization concentrate signal energy into a small number of coefficients, allowing low-energy components to be quantized more aggressively.

The quality of the reconstructed image is typically evaluated using distortion metrics defined in the pixel domain. The most common metrics include Mean Squared Error (MSE), Peak Signal-to-Noise Ratio (PSNR), and the Structural Similarity Index (SSIM) (Wang et al., 2004).

Mean Squared Error measures the average squared difference between the original image X and the reconstructed image X̂:

$$\mathrm{MSE}=\frac{1}{N}\sum_{i=1}^{N}\left(x_i-\hat{x}_i\right)^2 \tag{2.5}$$

Peak Signal-to-Noise Ratio expresses this error relative to the maximum representable pixel value:

$$\mathrm{PSNR}=10\log_{10}\frac{\mathrm{MAX}^2}{\mathrm{MSE}} \tag{2.6}$$

where MAX denotes the maximum possible pixel value.

The Structural Similarity Index (SSIM) was introduced to incorporate perceptually meaningful comparisons of luminance, contrast, and structural information (Wang et al., 2004). Although SSIM represents a significant improvement over purely error-based measures such as MSE, it remains fundamentally a pixel-domain metric, evaluating similarity based on local intensity statistics rather than the semantic or clinical meaning of image structures.

The theoretical elegance of the Shannon–Berger framework lies in its abstraction: once the distortion measure is defined, compression becomes a well-posed optimization problem governed by mathematically precise limits. This abstraction has proven extraordinarily successful in natural image coding and underlies most modern compression standards.

However, the framework implicitly assumes that the chosen distortion metric adequately captures the information that must be preserved in the signal. In practice, this assumption links statistical properties of the image—particularly variance—to the perceived or functional importance of image structures. The next section examines this assumption more closely and explores its implications in the context of medical imaging.

2.3 The Silent Assumption: Variance as Information

Within the classical rate–distortion framework introduced in the previous section, compression performance is governed by the relationship between bitrate and reconstruction error. For the widely studied case of a Gaussian source with squared-error distortion, the rate–distortion function takes the form $R(D) = \frac{1}{2} log_2(\sigma^2 / D)$, where σ^2 denotes the variance of the source and D represents the expected distortion (Cover & Thomas, 2006).

This result reveals a fundamental property of classical compression theory: the bitrate required to represent a signal is directly related to its variance. Higher-variance signals require more bits to achieve a given distortion level, whereas lower-variance signals can be represented with fewer bits.

In practical compression systems, this relationship manifests through bit allocation strategies that prioritize signal components with larger variance or higher energy. Transform coding methods—such as discrete cosine transform (DCT), wavelet coding, and vector quantization—concentrate signal energy into a small set of coefficients. During quantization, high-energy coefficients are preserved with higher precision, while low-energy coefficients are quantized more aggressively or discarded.

From the perspective of rate–distortion optimization, this strategy is mathematically optimal: minimizing mean squared error naturally leads to allocating more bits to components that contribute most to the signal's variance.

Yet this optimization implicitly introduces a conceptual identification that is rarely stated explicitly:

$$\textit{pixel variance} \approx \textit{information importance} \tag{2.7}$$

This identification does not appear as an explicit assumption in the formal mathematics of rate–distortion theory. Rather, it emerges indirectly through the choice of distortion metrics such as MSE, PSNR, and related energy-based measures. As noted in classical treatments of compression, squared-error distortion is favored primarily because of its mathematical tractability rather than because it universally reflects perceptual or semantic importance (Sayood, 2017).

For natural images, this approximation often works reasonably well. Regions containing edges, textures, or objects of visual interest tend to exhibit higher local variance than uniform background areas, and preserving high-variance components typically preserves visually salient structures.

However, the validity of this approximation becomes questionable in domains where informational relevance does not correlate with variance.

2.4 Failure of the Variance Proxy in Medical Imaging

Medical imaging provides a particularly striking example of this mismatch. In many clinical imaging scenarios, diagnostically significant structures occupy only a small fraction of the image domain and contribute minimally to the global variance of the image. Yet these structures often carry the most important diagnostic information.

Consider several representative examples in thoracic CT imaging:

Ground-glass nodules. Early-stage lung adenocarcinomas frequently appear as ground-glass opacities whose attenuation differs from surrounding lung parenchyma by only tens of Hounsfield units. Despite their subtle intensity contrast and small spatial footprint, their presence fundamentally alters clinical diagnosis and patient management.

Vascular bifurcations. The topology of the pulmonary vascular tree—including the presence or absence of bifurcations—can reflect perfusion abnormalities, tumor invasion, or vascular occlusion. These structures occupy

only a few pixels within a slice and contribute negligibly to the global variance of the image.

Tumor spiculation. Spiculated tumor margins are among the strongest radiological indicators of malignancy. These structures consist of thin, high-frequency protrusions extending from the tumor boundary. Because they occupy a very small number of pixels, their contribution to overall image variance is extremely small despite their diagnostic importance.

In each of these scenarios, the structures that determine clinical interpretation occupy a small fraction of the image variance but a large fraction of the diagnostic information. Classical compression frameworks, however, allocate representational resources proportionally to variance. Structures that contribute little to global variance receive fewer bits during encoding, making them particularly vulnerable to smoothing or removal during aggressive compression.

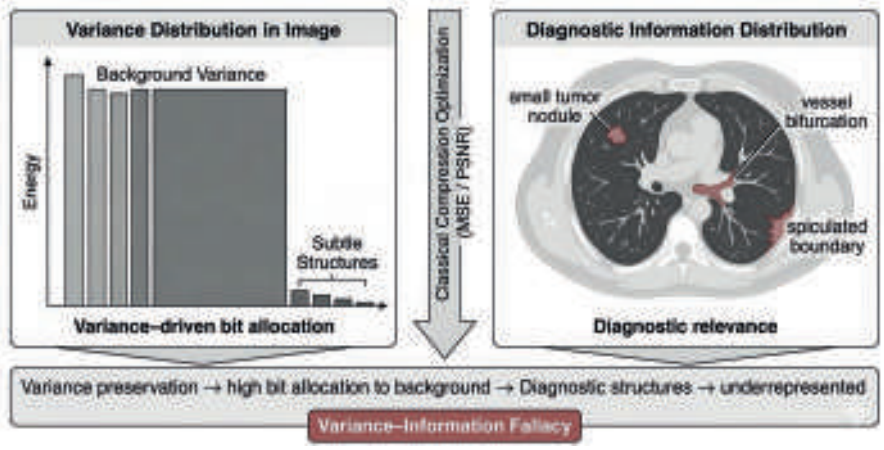

Figure 2.2 — Variance-driven compression versus diagnostic information distribution in medical images.
Classical compression algorithms allocate representational resources according to signal variance because distortion metrics such as MSE and PSNR are energy-based. In medical images, however, diagnostically significant structures—such as small nodules, vascular bifurcations, or tumor spiculation—occupy a very small fraction of the image domain and contribute little to global variance. Consequently, variance-driven bit allocation may preferentially preserve background variance while degrading structures that carry critical diagnostic information.

2.5 Formalizing the Variance–Information Fallacy

This discrepancy can be illustrated through a simplified model of a medical image. Let X denote a medical image and partition the pixel domain into

two subsets: R_d (pixels belonging to diagnostically relevant structures) and R_d (background pixels). Define the diagnostic variance ratio

$$\eta = \frac{\sum_{i \in R_d} \left(x_i - \mu \right)^2}{\sum_{i \in X} \left(x_i - \mu \right)^2} \tag{2.8}$$

where μ is the global mean intensity. For typical thoracic CT images, diagnostically relevant structures often occupy only a small portion of the image domain. In a simplified illustrative scenario where diagnostically relevant regions occupy approximately 5% of the image and exhibit variance comparable to the global average, the diagnostic variance ratio becomes $\eta \approx 0.04$, meaning that diagnostically relevant structures contribute roughly 4% of the total image variance.

Under an MSE-optimal compression strategy, bit allocation tends to follow variance distribution. Consequently, approximately 96% of the bit budget is devoted to preserving background variance, while only a small fraction of representational capacity is devoted to preserving diagnostically critical structures.

From the perspective of rate–distortion optimization, this allocation is mathematically optimal. Yet from the perspective of clinical information preservation, it is fundamentally misaligned with the diagnostic task.

This mismatch is what we term the Variance–Information Fallacy: the implicit assumption that preserving variance in the pixel domain ensures preservation of diagnostically relevant information.

It is important to emphasize that the fallacy does not arise from an error in the mathematics of information theory. Shannon's framework and the rate–distortion function remain mathematically correct. The problem arises when the variance-based distortion metrics used in practical compression systems are interpreted as proxies for informational relevance in domains where that relationship does not hold.

The next section examines this issue in a concrete algorithmic setting by analyzing the behavior of the SYMPES compression framework as a representative implementation of the classical paradigm.

2.6 The SYMPES Framework: A Classical Compression Reference

To examine the practical consequences of the variance–information assumption, a concrete compression framework is needed. The SYMPES

(Symmetric Pattern and Energy Synthesis) algorithm provides a suitable reference implementation because it operates within the classical rate–distortion paradigm and its internal mechanisms can be analyzed in terms of the theoretical concepts introduced in the preceding sections.

2.6.1 Block Decomposition and Signal Representation

The SYMPES framework models an image as a collection of local signal patches whose structure can be approximated through a small set of representative atoms. Instead of treating the image as a continuous field, the encoder partitions the image into non-overlapping square blocks of size $B \times B$ pixels. In the reference implementation, $B=8$, yielding a block dimensionality of $n=B^2=64$.

Let $x \in R^n$ denote the vectorized representation of a block extracted from the image. The fundamental assumption of the SYMPES representation is that the block structure can be approximated as the product of three components:

$$x \approx g\left(\mathfrak{e} \odot p\right) \tag{2.9}$$

where $\mathfrak{e}$ denotes an energy, atom selected from the Classified Energy Bank (CEB), p denotes a pattern atom selected from the Classified Pattern Bank (CPB), g is a scalar gain coefficient, and $\odot$ represents the element-wise (Hadamard) product.

The product $\left(\mathfrak{e} \odot p\right)$ produces a combined atom representing both the amplitude structure and the spatial arrangement of the block. This decomposition separates two aspects of local image structure: the energy distribution of the signal and the spatial pattern that organizes that energy. Such factorization allows the encoder to represent a large variety of local image configurations using a compact dictionary of atoms.

2.6.2 Joint Atom Selection and Gain Estimation

For each block x, the encoder performs a joint search over candidate energy–pattern atom pairs $\left(\mathfrak{e}_i, p_j\right)$ drawn from the two dictionaries. The objective is to identify the pair that best approximates the block under a squared-error criterion. The selection criterion can be expressed as the maximization of normalized correlation:

$$\left(i^*, j^*\right) = \arg\max \frac{\left(x^{\mathrm{T}}\left(\mathfrak{e}_i \odot p_j\right)\right)^2}{\| \mathfrak{e}_i \odot p_j \|^2} \tag{2.10}$$

This criterion is equivalent to maximizing the cosine similarity between the block and the candidate atom, which in turn minimizes the squared reconstruction error. Once the optimal atom pair has been selected, the scalar gain coefficient is obtained through least-squares projection:

$$g^* = \frac{x^{\mathrm{T}} \left(\mathfrak{e}_{i^*} \odot p_{j^*} \right)}{\| \mathfrak{e}_{i^*} \odot p_{j^*} \|^2} \tag{2.11}$$

The reconstructed block is then given by

$$\hat{x} = g^* \left(\mathfrak{e}_{i^*} \odot p_{j^*} \right) \tag{2.12}$$

The difference $r = x - \hat{x}$ represents the residual signal that cannot be captured by the selected atom pair. The squared residual norm $\|r\|^2$ contributes directly to the Mean Squared Error used to evaluate compression performance.

An important property of this formulation is that the optimization criterion is entirely driven by energy capture. The algorithm selects atoms that maximize the fraction of block variance explained by the representation, without distinguishing between diagnostically relevant and irrelevant components of the signal.

2.6.3 Dictionary Construction: CEB and CPB

The two dictionaries used in the SYMPES representation are constructed from training data through unsupervised clustering and redundancy pruning.

Classified Energy Bank (CEB). Training blocks are first normalized to unit ℓ_2-norm to remove global intensity scaling. The normalized vectors are then clustered using the K-means algorithm to obtain a set of representative centroids that capture dominant energy distributions within the dataset (Gersho & Gray, 1992; Goyal, 2001). Because K-means clustering minimizes within-cluster variance, the resulting centroids correspond to directions that capture the most prominent energy patterns present in the training data. To avoid redundancy in the dictionary, candidate atoms whose Pearson correlation exceeds a predefined threshold are removed through greedy sequential pruning.

Classified Pattern Bank (CPB). Once the energy atoms are determined, pattern atoms are extracted from residual structures that remain after energy normalization. For each training block x, the best matching energy atom e and gain g are first estimated. A normalized residual pattern is then computed

as $p = x / (g \cdot \epsilon + \varepsilon)$, where ε is a small stabilization constant preventing division by zero. These residual patterns are subsequently clustered using K-means in the same manner as the energy atoms. After redundancy pruning, the resulting centroids form the Classified Pattern Bank (CPB).

The combined use of CEB and CPB provides a compact representation capable of approximating a wide range of block configurations. However, the dictionary construction process reveals an important structural property of the model. Because both dictionaries are derived through variance-minimizing clustering, the atoms primarily represent statistically dominant structures in the training dataset. Rare but diagnostically significant structures—such as subtle tumor boundaries or fine vascular branches—may be poorly represented because they contribute little to the global variance distribution.

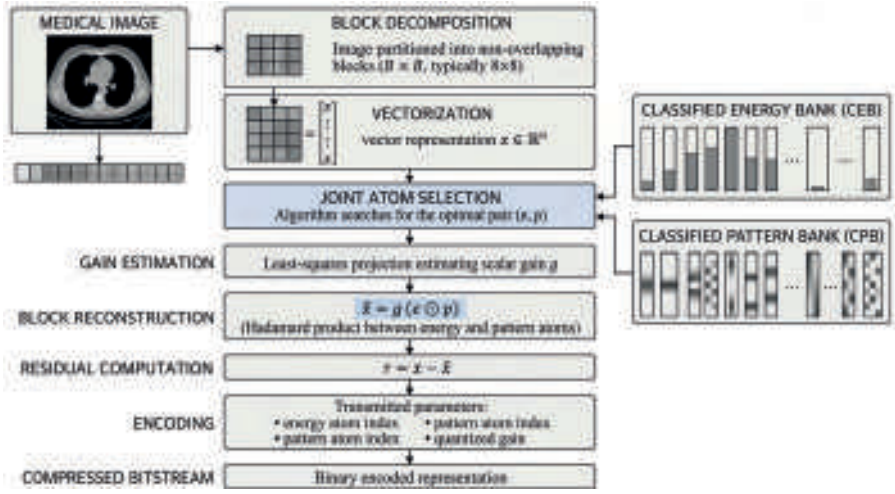

Figure 2.3 — Algorithmic structure of the SYMPES compression framework. The input medical image is partitioned into non-overlapping blocks that are vectorized and represented through a pair of atoms selected from two learned dictionaries: the Classified Energy Bank (CEB) and the Classified Pattern Bank (CPB). For each block, the encoder identifies the optimal atom pair that maximizes normalized correlation with the signal and estimates a scalar gain coefficient via least-squarestprojection. The reconstructed block is obtained as the gain-scaled Hadamard product of the selected atoms, and the encoder transmits the indices of the atoms together with the quantized gain parameter.

2.6.4 Compression Levels and Bitrate Representation

The SYMPES framework defines multiple compression levels by varying the dictionary sizes and gain quantization precision. For each block, the encoder transmits three elements: the index of the selected energy atom, the index of the selected pattern atom, and the quantized gain coefficient. The total number of bits required per block is therefore $b_{block} = log_2(M) + log_2(N) + q$, where M is the number of energy atoms, N is the number of pattern atoms, and q is the gain quantization precision.

The corresponding bits per pixel (BPP) is $BPP = b_{block} / B^2$, which determines the compression ratio relative to the original image bit depth.

From the perspective of classical rate–distortion theory, this representation behaves as expected: reducing the dictionary size or gain precision lowers the bitrate while increasing reconstruction error. The mathematical coherence of this framework illustrates how classical compression systems allocate representational resources in proportion to signal variance. Yet the optimization objective remains entirely agnostic to the semantic or diagnostic relevance of image structures.

The theoretical analysis presented in this chapter demonstrates that classical compression frameworks, while mathematically optimal under their stated assumptions, operate under an implicit identification of variance with information. In medical imaging, this identification fails precisely where diagnostic relevance is highest. Part III subjects this theoretical prediction to systematic empirical testing using clinical CT data from the TCGA-LUAD cohort.

Chapter Summary and Transition

This chapter established that classical compression theory is fundamentally grounded in variance-weighted distortion minimisation. Within the Shannon rate–distortion framework, variance is implicitly treated as a proxy for information, leading to the assumption that minimising pixel-domain error preserves signal content. The SYMPES framework was introduced as an analytically transparent reference implementation of this classical paradigm, illustrating how block decomposition, dictionary construction, and gain quantisation allocate representational resources in proportion to signal variance.

However, the analysis demonstrated that this assumption constitutes a structural oversimplification in the context of medical imaging. Clinical information is not uniformly distributed across the intensity space but is instead concentrated in sparse, anatomically meaningful structures—lesion boundaries,

vascular networks, parenchymal textures—and topological configurations that occupy only a small fraction of the global variance budget. The Variance–Information Fallacy, formalised in this chapter, describes this fundamental mismatch: variance preservation does not guarantee the preservation of diagnostically relevant information.

While this conclusion is theoretically grounded, it remains insufficient without empirical validation. The mathematical coherence of the classical framework does not, by itself, determine whether the predicted divergence between variance fidelity and diagnostic information manifests in real clinical imaging data. Specifically, the central question remains unresolved:

Does variance-optimised compression empirically preserve clinical structure in real-world medical imaging data?

To address this, the next chapter transitions from theoretical critique to empirical analysis. Using the TCGA-LUAD cohort as a heterogeneous clinical stress test, Part III investigates whether compression behaviour that appears stable under pixel-domain metrics remains stable under structurally and clinically grounded evaluation criteria. The multi-layer evaluation framework introduced therein—encompassing global pixel fidelity, region-specific structural fidelity, and topological stability—provides the empirical instruments necessary to test the predictions of the Variance–Information Fallacy under realistic clinical conditions.

CHAPTER 3

Empirical Analysis of Medical Image Compression

Chapter Thesis

Empirical analysis reveals a structural decoupling between pixel fidelity and clinically meaningful information in compressed medical images.

3.1 The Empirical Question: Does Variance-Weighted Compression Preserve Clinical Structure?

The theoretical arguments developed in Parts I and II culminate in a precise empirical question: does variance-weighted compression preserve clinically meaningful structure in medical images? Classical compression theory, when instantiated through squared-error distortion measures, allocates representational resources in proportion to intensity variance. Under this logic, the dominant objective of the encoder is not to preserve diagnostic salience, anatomical topology, or feature-level invariance, but to minimize average reconstruction error in the image domain. This principle is mathematically coherent within the Shannon–Berger rate–distortion framework (Shannon, 1948; Berger, 1971). The unresolved issue is whether that coherence survives contact with real clinical imaging data.

In natural-image settings, the implicit alignment between variance and informational relevance often remains serviceable because visually salient structures frequently coincide with high-energy image components. Medical imaging, however, is governed by a different geometry of information. Diagnostically important structures are often spatially sparse, morphologically subtle, and statistically underrepresented relative to the total image field. A small lesion boundary, a faint ground-glass opacity, a thin spiculation, or a

minor disruption in vascular continuity may occupy only a tiny fraction of the global variance budget while carrying disproportionate clinical meaning. Consequently, a compression system that is globally optimal in the rate–distortion sense may remain locally destructive in exactly those regions that matter most for diagnosis and downstream clinical inference.

This chapter therefore shifts the monograph from theoretical critique to empirical demonstration. The goal is not merely to show that compression changes images—that would be trivial—but to determine how those changes are distributed across levels of representation, and whether clinically meaningful degradation can emerge while conventional pixel-domain metrics remain apparently acceptable. In other words, the chapter tests the central claim of this monograph under realistic clinical conditions: pixel fidelity is a necessary but insufficient condition for diagnostic information preservation (Clunie, 2000; Zhu et al., 2018).

The empirical setting is the TCGA-LUAD cohort from The Cancer Imaging Archive (TCIA) (Clark et al., 2013), a multi-institutional thoracic CT collection that contains substantial heterogeneity in scanner manufacturer, acquisition settings, reconstruction behavior, and patient anatomy. This heterogeneity is not treated here as an inconvenient nuisance to be averaged away. On the contrary, it is analytically essential. If variance-preserving compression remains clinically meaningful only in artificially homogenized datasets, then its practical relevance to real-world imaging pipelines is already questionable. A clinically credible empirical test must therefore operate in a setting where biological signal and technical variance coexist and, in part, overlap. The TCGA-LUAD cohort provides exactly such a stress environment.

A second design principle of the chapter is methodological determinism. All analyzes are carried out through a fixed experimental pipeline in which pre-processing, compression, region definition, structural descriptor extraction, and statistical testing are applied identically across all compression levels. This is not a cosmetic implementation detail; it is a scientific necessity. Without deterministic control of the analysis chain, observed changes could be attributed to procedural drift rather than to compression itself. Since the central thesis concerns subtle decouplings between intensity-domain similarity and structural stability, eliminating avoidable methodological variance is essential.

The empirical logic of the chapter is organized around a three-layer evaluation framework (Figure 3.1). The experimental pipeline begins with cohort construction and standardized pre-processing, followed by deterministic compression and reconstruction across predefined operating levels. The resulting images are then evaluated at three complementary levels of representation.

The first layer examines global pixel-domain fidelity, using classical reconstruction metrics such as mean squared error (MSE), peak signal-to-noise ratio (PSNR), and structural similarity (SSIM). This layer addresses the conventional question posed by compression literature (Aiazzi et al., 2003; Bilgin et al., 2001): how closely does the reconstructed image resemble the original when averaged across the full image domain?

The second layer evaluates ROI-specific fidelity, restricting the analysis to lung parenchyma and lesion-bearing regions where diagnostically relevant structures are concentrated. By isolating anatomically meaningful regions of interest, this layer tests whether global averages systematically obscure localized degradation within clinically critical structures.

The third layer investigates higher-order representational stability, including structural descriptor drift and topological change under increasing compression. Even when pixel-level similarity remains high, compression may perturb spatial organization, texture statistics, or topological connectivity in ways that affect radiological interpretation or downstream artificial-intelligence models. This layer therefore probes a deeper and more clinically consequential question: whether compressed images preserve the structural organization upon which diagnostic reasoning ultimately depends.

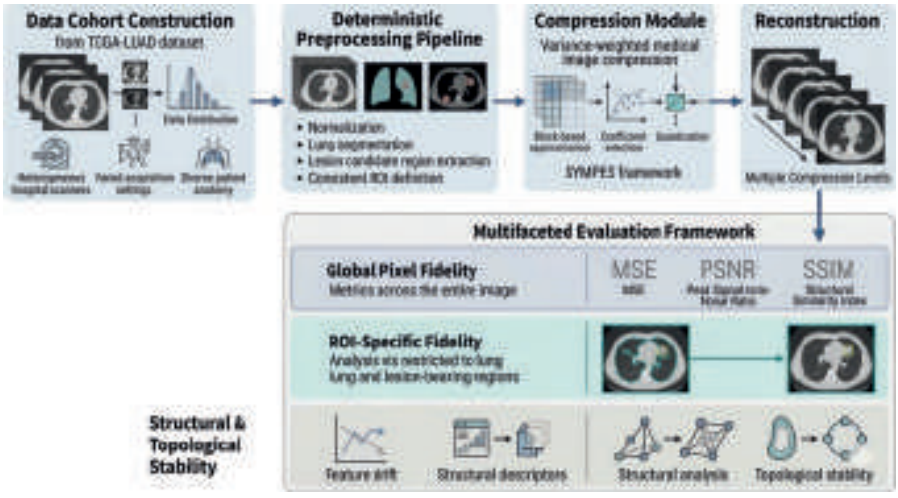

Figure 3.1. Deterministic experimental design and three-layer evaluation framework for assessing the effects of variance-weighted medical image compression in heterogeneous thoracic CT data. The pipeline proceeds from cohort construction and standardised preprocessing to controlled compression and reconstruction. Compression effects are evaluated at three complementary levels: global pixel fidelity (MSE, PSNR, SSIM), ROI-specific fidelity within lung and lesion regions, and higher-order structural and topological stability.

As illustrated in Figure 3.1, these three evaluation layers operate on the same deterministic experimental pipeline. The identical pre-processing procedures, compression configurations, and ROI definitions are applied across all compression levels. This controlled structure ensures that any observed degradation arises from the compression process itself rather than from variability in the analysis procedure.

The role of the Classical SYMPES framework in this chapter is correspondingly precise. SYMPES is not introduced here as a universal solution or as a modern benchmark intended to outperform alternative compression methods. Instead, it functions as an analytically transparent reference implementation of the classical paradigm. Because its block structure, atom selection mechanism, gain quantization strategy, and variance-driven approximation behavior are explicitly defined, it provides a concrete operational environment in which the theoretical assumptions of Part II can be examined empirically. In this sense, the original doctoral framework serves not only as historical precedent but also as a methodological bridge connecting classical rate–distortion reasoning with observable behavior in real clinical CT data.

The chapter proceeds in stages from dataset construction to progressively deeper levels of representational risk. After defining the cohort and its heterogeneity profile, the deterministic evaluation pipeline is formalized and applied to reconstructed images generated across multiple compression levels. Global pixel-level behavior is examined first, followed by ROI-specific degradation patterns and finally by structural and topological stability analyzes. This staged progression is deliberate. It allows the reader to observe how a compression process may appear well behaved under conventional fidelity criteria while simultaneously failing in clinically meaningful representational subspaces.

While classical compression theory evaluates fidelity through distortion measures such as mean squared error, these measures implicitly assume that variance preservation is an adequate surrogate for informational preservation (Shannon, 1948; Berger, 1971). In medical imaging, however, diagnostically relevant structures often occupy only a small fraction of the global variance budget (Clunie, 2000). This structural mismatch between variance-driven optimization and diagnostic salience constitutes the central empirical question of the present chapter.

Ultimately, this chapter does not attempt to reject classical compression theory on its own mathematical terms. The principles of rate–distortion optimization remain internally consistent and theoretically elegant. The problem lies elsewhere: in the implicit assumption that distortion measures

optimized for variance preservation are sufficient surrogates for the preservation of clinical information. The empirical analyzes that follow demonstrate that, in heterogeneous thoracic CT data, this assumption can fail quietly, systematically, and in diagnostically important ways.

3.2 TCGA-LUAD as a Heterogeneous Clinical Stress Test

A meaningful empirical evaluation of medical image compression requires more than data availability; it requires a dataset whose internal structure exposes the tension between statistical fidelity and clinical relevance. For the purposes of this chapter, the TCGA-LUAD cohort from The Cancer Imaging Archive (TCIA) was selected because it provides a clinically authentic, multi-institutional, and technically heterogeneous thoracic CT environment in which biological variability and scanner-dependent variance coexist. Such characteristics make the cohort particularly suitable for testing whether compression behavior that appears acceptable under conventional pixel-domain metrics remains acceptable under clinically structured evaluation. The TCIA repository provides curated public access to cancer imaging datasets collected from multiple institutions and scanner platforms, enabling reproducible quantitative imaging research (Clark et al., 2013).

The empirical role of this dataset must also be interpreted in relation to the deterministic experimental pipeline defined in Chapter 3.1. As illustrated in Figure 3.1, all images entering the evaluation framework pass through a controlled sequence of pre-processing, compression, reconstruction, and multi-layer fidelity analysis. The dataset therefore functions not merely as a collection of images but as the input domain for a deterministic experimental system designed to probe the relationship between variance preservation and diagnostic information stability.

The raw TCGA-LUAD archive contains multiple imaging modalities and multiple imaging series per patient. Let P denote the set of patients in the raw archive, and let $S(p)$ denote the set of imaging series associated with patient $p \in P$. Because a single patient may contribute several series of differing modality, acquisition parameters, or completeness, a deterministic series-selection rule must be defined before any compression analysis can be meaningfully performed. Accordingly, a selection operator Φ was introduced to select a single CT series per patient subject to explicit constraints on modality compatibility, spatial resolution, and volumetric adequacy. Formally, the selected series s_p^* for patient p is defined as:

$$s_p^* = \Phi\left\{S(p)\right\} = arg \max_{s \hat{\partial} S(p)} N_{slices}(s) \tag{3.1}$$

This formulation enforces three methodological principles simultaneously. First, it restricts the empirical analysis to CT acquisitions, thereby ensuring modality coherence across the cohort. Second, it removes incompatible in-plane resolutions that would otherwise introduce artificial variability during pre-processing. Third, it favors the volumetrically richest CT series available for each patient, preserving maximal anatomical continuity for subsequent ROI-based and structural analyzes. The filtering procedure therefore functions as a methodological safeguard against confounding variability introduced prior to compression.

The transition from raw archive to analytic cohort is summarized in Figure 3.2, which presents representative axial slices from multiple patients after deterministic filtering. The figure illustrates the anatomical diversity of the cohort, including variations in lung morphology, lesion appearance, and surrounding thoracic structures. Such inter-patient variability reflects the biological diversity of lung cancer imaging and provides the anatomical context within which compression behavior must be evaluated.

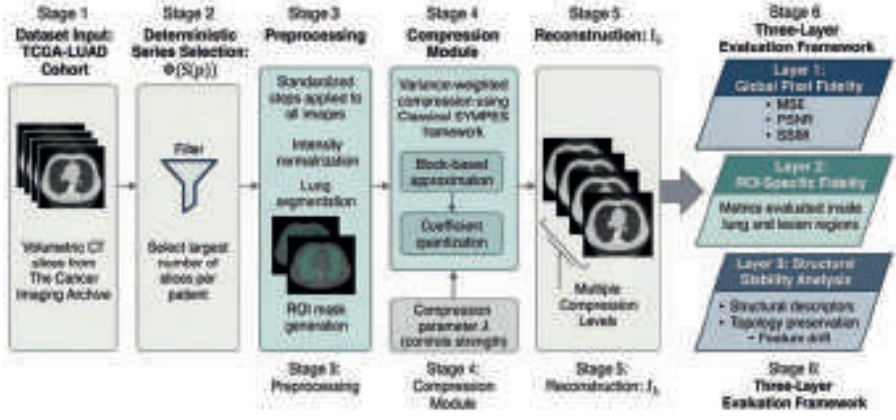

Figure 3.2. Representative axial CT slices from the TCGA-LUAD cohort after deterministic filtering. Each row corresponds to a different patient, illustrating anatomical variability and lesion diversity across the dataset.

The raw archive originally contained 69 subjects and 624 imaging series spanning several modalities, including 551 CT series, 66 PT series, and 7 NM series, corresponding to an average of 9.04 series per subject. After deterministic filtering, the final analytic cohort consisted of 67 patients, 69 CT series, and 335 axial slices selected for detailed compression analysis. Series were excluded for three principal reasons: insufficient slice count, non-CT

modality, or resolution mismatch. These exclusion statistics document the transition from a heterogeneous clinical archive to a compression-ready dataset whose variability remains clinically realistic but methodologically tractable.

The scientific relevance of this cohort lies not simply in its size but also in its heterogeneity profile at the scanner level. The retained CT series originate from multiple scanner manufacturers, including Siemens, GE Medical Systems, CPS, ADAC, and CTI. The distribution of manufacturers is shown in Figure 3.3. Multi-vendor acquisition environments introduce structured technical variance into reconstructed CT images through differences in detector architecture, reconstruction kernels, proprietary filtering pipelines, and slice-thickness conventions. Such scanner-dependent factors interact with biological signal during image reconstruction, producing intensity and texture patterns that cannot be trivially separated into purely biological or purely technical components.

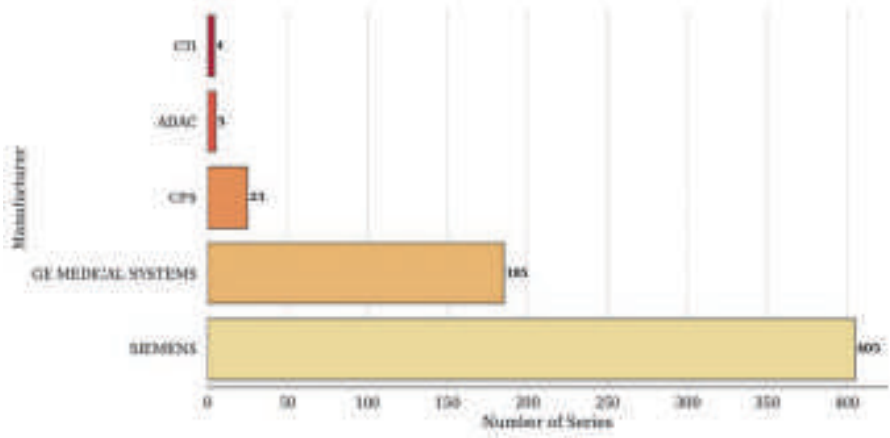

Figure 3.3. Scanner manufacturer distribution in the TCGA-LUAD cohort, including Siemens, GE Medical Systems, CPS, ADAC, and CTI platforms. The multi-vendor composition reflects the technical heterogeneity typical of multi-institutional clinical imaging archives.

The presence of this heterogeneity is not a limitation but an empirical advantage. In homogeneous imaging datasets, scanner-induced variance can be artificially minimized, producing an environment in which variance-preserving compression may appear deceptively stable. In contrast, multi-institutional datasets such as TCGA-LUAD preserve the technical diversity characteristic of real clinical imaging environments. Evaluating compression within this

context therefore provides a more realistic assessment of whether variance-weighted reconstruction truly preserves diagnostically relevant structures.

Another important property of the cohort concerns its volumetric variability. The slice-count distribution across retained CT series is markedly right-skewed, reflecting differences in acquisition coverage and reconstruction protocols across institutions. This distribution is shown in Figure 3.4. Rather than clustering around a narrowly standardised acquisition depth, the cohort spans a broad range of volumetric extents, indicating that patients are represented by CT series of differing anatomical coverage and through-plane continuity. This heterogeneity is methodologically important because thoracic CT interpretation depends not only on the fidelity of isolated axial images but also on the preservation of relationships across adjacent slices. Lung boundaries, vascular trajectories, lesion morphology, and contextual tissue relationships unfold across contiguous sections; therefore, compression effects cannot be interpreted solely at the level of single-slice reconstruction quality.

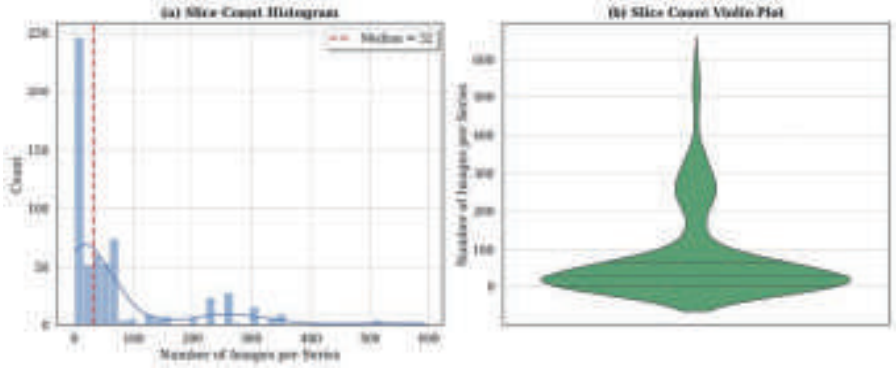

Figure 3.4. Distribution of slice counts across the retained CT series in the TCGA-LUAD cohort. The right-skewed profile reflects volumetric heterogeneity in acquisition depth and anatomical coverage across patients.

The scientific relevance of this volumetric heterogeneity is twofold. First, it demonstrates that the selected dataset is not artificially homogenised around a fixed acquisition template. Second, it creates a more demanding evaluation regime for compression, since methods that appear stable in isolated or low-depth images may behave differently when anatomical information is distributed across longer volumetric sequences. In this respect, Figure 3.3 does not merely describe the dataset; it supports the central experimental claim that clinically meaningful compression analysis must account for the continuity and structural dependence of volumetric thoracic CT data.

In this sense, the TCGA-LUAD dataset functions as an empirical analog of the theoretical asymmetry described in Part I of this monograph. There, reconstructed CT variance was shown to incorporate not only biological heterogeneity but also scanner-induced technical signatures and measurement-driven structural effects. Within a heterogeneous clinical archive, these components coexist in the reconstructed intensity field. Consequently, a compression system that preserves dominant variance components may still perturb diagnostically meaningful structures embedded within this mixed variance landscape.

For the purposes of this chapter, the TCGA-LUAD cohort should therefore be interpreted not merely as a dataset but as a clinical stress environment for variance-weighted compression. Figure 3.2 establishes the anatomical diversity of the retained cohort, Figure 3.3 documents its scanner-level technical diversity, and Figure 3.4. demonstrates its volumetric heterogeneity. Taken together, these three figures define the empirical conditions under which the central hypothesis of this chapter can be tested. If compression algorithms are genuinely aligned with clinically meaningful information, their behavior should remain stable despite the biological and technical heterogeneity present in the cohort. If, however, compression merely preserves dominant variance patterns without regard to structural meaning, then a heterogeneous dataset such as TCGA-LUAD provides precisely the conditions under which that failure becomes observable. This empirical tension between statistical fidelity and structural stability forms the central experimental premise of the analyzes that follow.

Taken together, Figures 3.2–3.4 demonstrate that the TCGA-LUAD cohort embodies three interacting forms of heterogeneity: anatomical diversity across patients, volumetric variability across imaging series, and technical variance introduced by multi-vendor acquisition environments. These properties transform the dataset from a simple image collection into a clinically realistic stress environment for compression analysis. In such an environment, compression behavior cannot be interpreted solely through global pixel-domain fidelity metrics, because anatomical continuity, structural context, and scanner-dependent variance jointly influence how diagnostic information is represented within the reconstructed image field.

Establishing this empirical environment is a necessary prerequisite for the analyzes that follow. However, demonstrating the structural diversity of the dataset alone is not sufficient; a formal experimental framework must also be defined to ensure that compression behavior can be evaluated in a controlled and reproducible manner. Accordingly, the next section introduces

the experimental protocol used throughout this chapter. Section 3.3 specifies the compression algorithms under evaluation, the reconstruction procedures applied to the compressed images, and the quantitative fidelity metrics used to assess reconstruction behavior. By integrating these components into the deterministic pipeline introduced in Section 3.1 and the heterogeneous dataset environment characterized in Section 3.2, the experimental design establishes a structured framework for examining the relationship between variance preservation and diagnostic information stability in thoracic CT compression.

3.3 Compression Experiment and Multi-Layer Structural Evaluation

Following the deterministic dataset preparation described in Sections 3.1 and 3.2, the next step is to establish an experimental protocol capable of evaluating compression behavior within a clinically meaningful representational framework. Conventional compression studies primarily rely on global pixel-domain fidelity metrics such as mean squared error (MSE), peak signal-to-noise ratio (PSNR), and structural similarity index (SSIM). While these measures quantify reconstruction accuracy in a statistical sense, they implicitly assume that preserving global variance patterns is equivalent to preserving diagnostically relevant image information.

In medical imaging, however, clinically meaningful structures often occupy only a small fraction of the image domain. Lesion boundaries, vascular structures, and subtle parenchymal textures represent localized patterns whose statistical contribution to global variance may be minimal. Consequently, compression algorithms optimized for variance preservation may maintain high global fidelity while perturbing anatomically meaningful structures. This limitation has been widely discussed in radiomics and quantitative imaging literature, where the stability of image-derived features depends strongly on acquisition parameters and reconstruction characteristics (Aerts et al., 2014; Lambin et al., 2012; Gillies et al., 2016).

To address this limitation, the present study adopts a multi-layer evaluation framework that integrates pixel-domain fidelity analysis, region-specific structural evaluation, and topological stability assessment. The experimental pipeline therefore extends beyond traditional reconstruction metrics and examines compression behavior across multiple levels of image representation.

The algorithmic architecture of the SYMPES framework—including its block decomposition, dictionary construction (CEB and CPB), and multi-level quantization pipeline—has been described in detail in Chapter 2 (see Figure 2.3 and Section 2.6). The present chapter takes that architecture as given and

evaluates its empirical behavior under the specific conditions of 4D CT lung imaging. Accordingly, the analysis below proceeds directly to the structural fidelity metrics that govern the evaluation protocol.

3.3.1 Lung Region-of-Interest Extraction

To enable region-specific fidelity analysis, lung regions were extracted using automated segmentation. The resulting lung masks are illustrated in Figure 3.5, where green contours indicate the boundaries of the lung parenchyma used for ROI-based analysis.

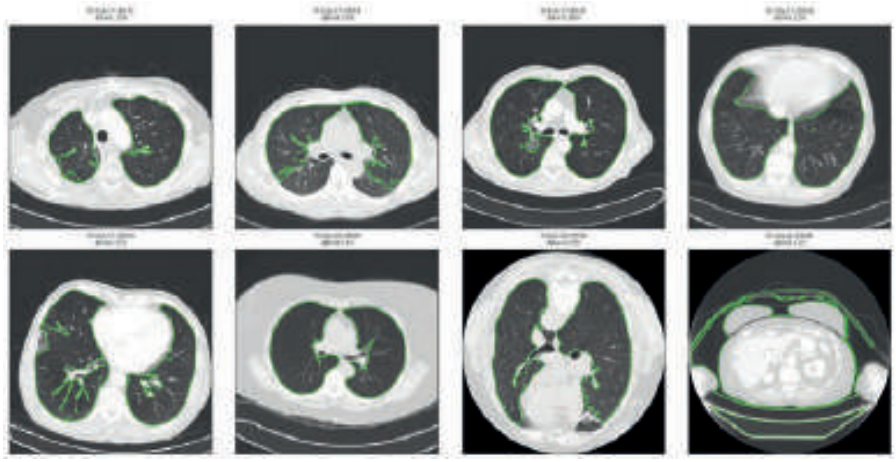

Figure 3.5. Automated lung segmentation results used to define the region-of-interest (ROI) for compression evaluation. Green contours represent the extracted lung boundaries.

The segmentation step allows compression effects to be evaluated separately within the lung region and outside it. This distinction is essential because background regions such as air or soft tissue outside the lungs may dominate global fidelity metrics despite having little clinical relevance.

3.3.2 Lesion Candidate Identification

Within the lung ROI, potential lesion regions were further identified using morphological filtering and intensity-based candidate detection. Example lesion candidates are shown in Figure 3.6, where red markers highlight localized abnormal structures.

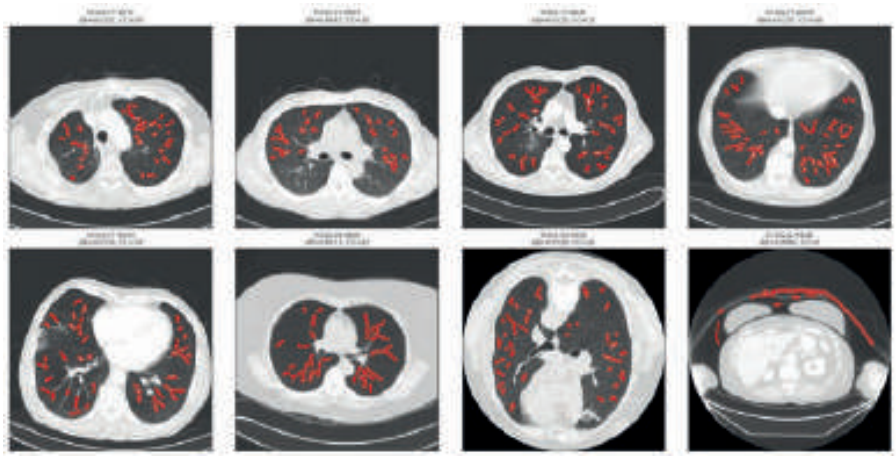

Figure 3.6. Lesion candidate localization within lung CT images. Red markers indicate regions used for localized compression fidelity analysis.

By isolating these candidate regions, compression behavior can be evaluated at the level of clinically relevant structures rather than across the entire image domain.

3.3.3 ROI Area Distribution

The statistical distribution of lung and lesion ROI areas across the dataset is summarized in Figure 3.7. The results reveal a strong asymmetry between the dominant anatomical regions and localized lesion structures.

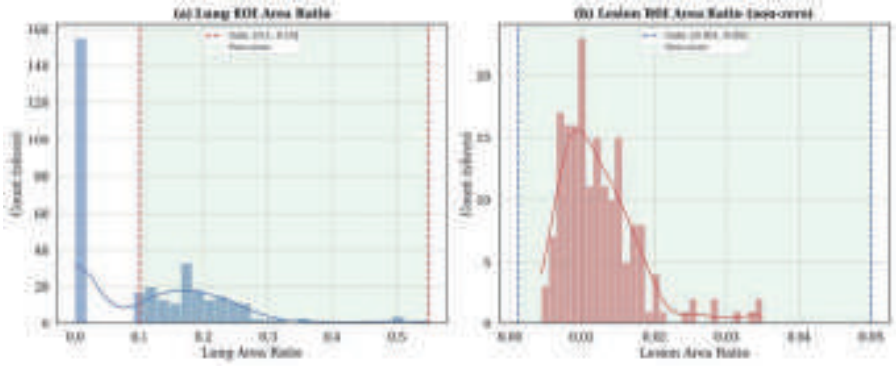

Figure 3.7. Distribution of lung ROI area ratio and lesion ROI area ratio across CT slices.

The lung ROI typically occupies between approximately 10% and 50% of the image domain, while lesion ROIs represent only a small fraction of the slice area. This imbalance highlights a key limitation of global compression metrics: preserving variance in the dominant background regions may not guarantee preservation of small but diagnostically critical structures.

3.3.4 Global vs ROI Fidelity Analysis

Compression experiments were performed across multiple compression levels ranging from mild compression (Q1) to aggressive compression (Q5). Global fidelity metrics were computed across the entire image domain, while ROI-specific metrics were computed separately for lung and lesion regions. The resulting PSNR and SSIM curves are presented in Figure 3.8.

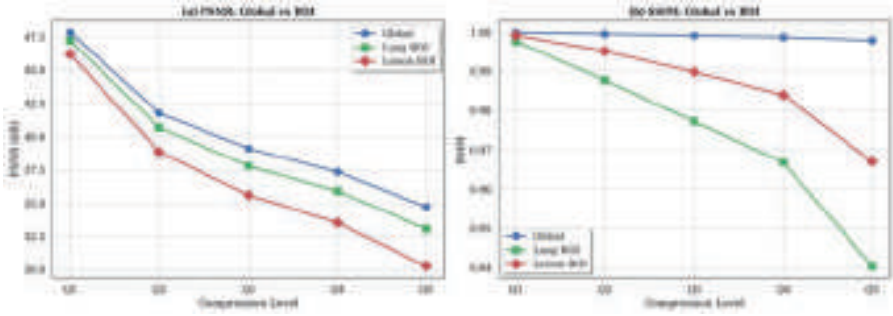

Figure 3.8. Comparison of global and ROI-specific fidelity metrics under increasing compression levels.

The results show that global PSNR and SSIM values remain relatively stable across moderate compression levels. However, ROI-specific metrics exhibit a more pronounced decline, particularly within lesion regions. This divergence suggests that compression artifacts disproportionately affect localized structures even when global reconstruction metrics remain favorable.

The difference between background and ROI fidelity is further quantified in Figure 3.9, which plots the PSNR gap between ROI regions and background pixels.

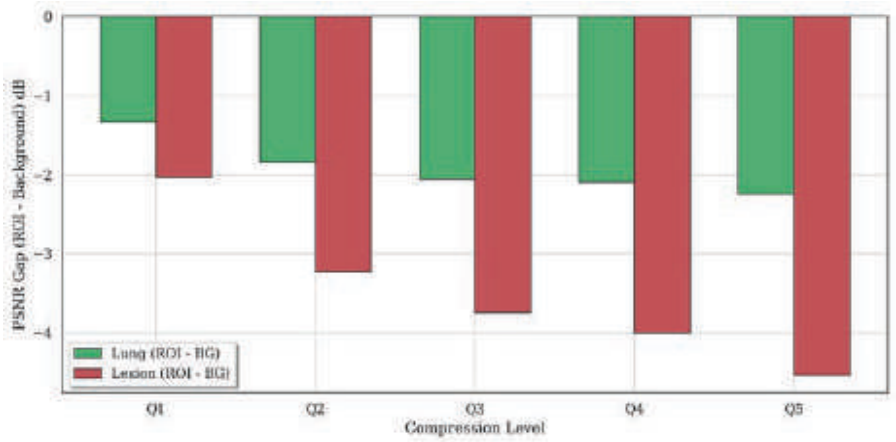

Figure 3.9. PSNR gap between ROI and background regions across compression levels.

The increasing gap demonstrates that compression errors accumulate more rapidly within clinically relevant regions than within background areas. This observation supports the hypothesis that variance-based compression objectives may favor background regions while degrading diagnostically meaningful structures.

3.3.5 Fidelity–Topology Decoupling

Beyond intensity-based fidelity measures, compression effects were also evaluated using topological descriptors of anatomical structure. Persistent homology analysis was used to quantify structural changes in lung morphology under compression. The relationship between PSNR decline and topological instability is illustrated in Figure 3.10.

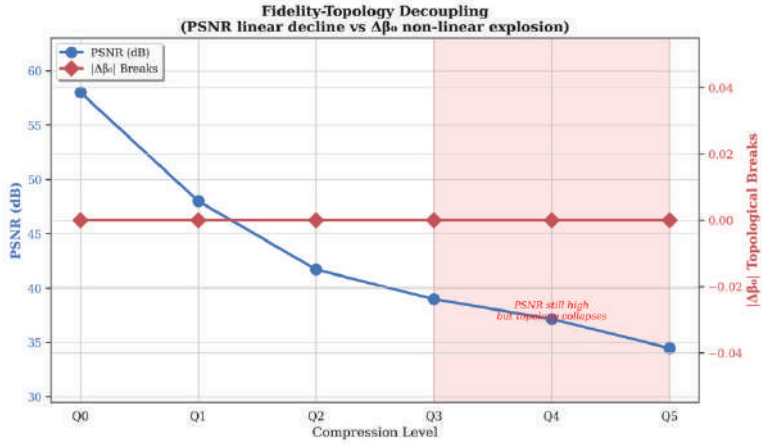

Figure 3.10. Fidelity–topology decoupling under increasing compression levels. While PSNR decreases approximately linearly, topological break events increase nonlinearly.

This behavior reveals a fundamental asymmetry between statistical fidelity and structural stability. Although global reconstruction error increases gradually with compression strength, structural break events occur abruptly once compression exceeds a critical threshold. The quantitative feature stability under compression is summarized in Table 3.1.

Table 3.1. Feature stability under compression. Similarity metrics between features extracted from original and compressed CT images. High stability indicates that compression does not significantly perturb quantitative descriptors used in radiomics pipelines.

Feature	Similarity Metric	Interpretation
Intensity histogram	Pearson correlation	High stability indicates minimal histogram perturbation
GLCM texture	Concordance correlation	Sensitive to spatial texture disruption
Topological features	Bottleneck distance	Captures structural connectivity changes

3.3.6 Phase Transition in Structural Stability

The structural consequences of compression can be further understood as a phase transition phenomenon. Segmentation overlap metrics remain relatively stable under mild compression but deteriorate rapidly once compression passes

a critical regime. This transition is shown in Figure 3.11, where Dice similarity and topological distance are plotted across compression levels.

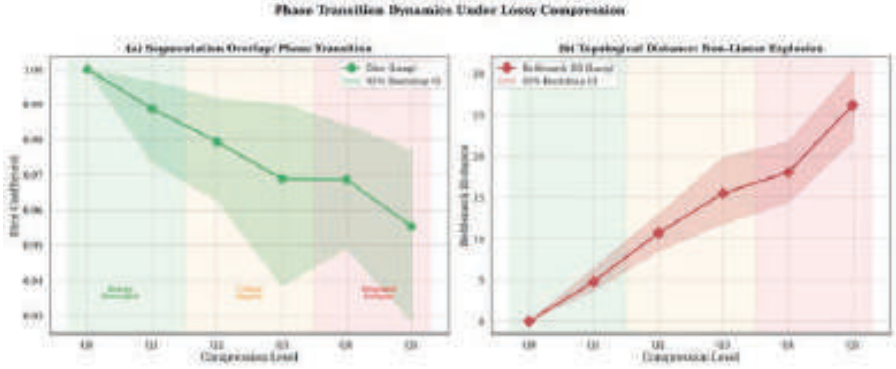

Figure 3.11. Phase transition dynamics in structural fidelity under lossy compression.

Three regimes can be observed: an energy-dominated regime at low compression with minimal structural distortion; a critical regime at intermediate compression where structural degradation begins; and a structural collapse regime at high compression where anatomical topology becomes unstable. Such behavior indicates that compression effects on anatomical structure are inherently nonlinear. The ROI-based fidelity metrics across these regimes are summarized in Table 3.2.

Table 3.2. ROI-based fidelity metrics. PSNR and SSIM values computed separately for global image regions, lung ROI regions, and lesion ROI regions across compression levels.

Region	PSNR (dB)	SSIM	Interpretation
Global image	Highest	Highest	Dominated by background
Lung ROI	Lower	Lower	Parenchyma degradation visible
Lesion ROI	Lowest	Lowest	Diagnostic structures most affected

The hierarchical evaluation framework linking compression behavior to progressively richer representations of image information is formalised in Table 3.3.

Table 3.3. Three-layer evaluation framework. This hierarchical evaluation framework links compression behavior to progressively richer representations of image information.

Layer	Representation	Evaluation Metric
Pixel layer	Global intensity field	PSNR, SSIM
Feature layer	Radiomic descriptors	Feature correlation
Structural layer	Anatomical topology	Persistent homology

3.4 Empirical Behavior of Compression under Structural Evaluation

The evaluation framework introduced in Section 3.3 enables the analysis of compression behavior beyond conventional pixel-domain fidelity metrics. Whilst classical compression evaluation relies primarily on global distortion measures such as peak signal-to-noise ratio (PSNR) and structural similarity index (SSIM), these metrics alone cannot capture the preservation of diagnostically meaningful structures. Medical images contain clinically relevant information concentrated in sparse anatomical regions—lesion boundaries, vascular structures, and texture variations—that occupy only a small fraction of the overall image domain (Lambin et al., 2012; Gillies et al., 2016).

To investigate how compression affects these structures, the present analysis evaluates compression behavior across three complementary layers: global pixel fidelity, region-specific reconstruction fidelity, and structural/topological stability. The following sections analyze how these layers evolve as compression strength increases.

3.4.1 Global Pixel-Domain Fidelity Behavior

Global compression behavior was first analyzed using conventional rate–distortion metrics. Figure 3.12 presents the rate–distortion trajectory of the SYMPES compression framework across compression levels.

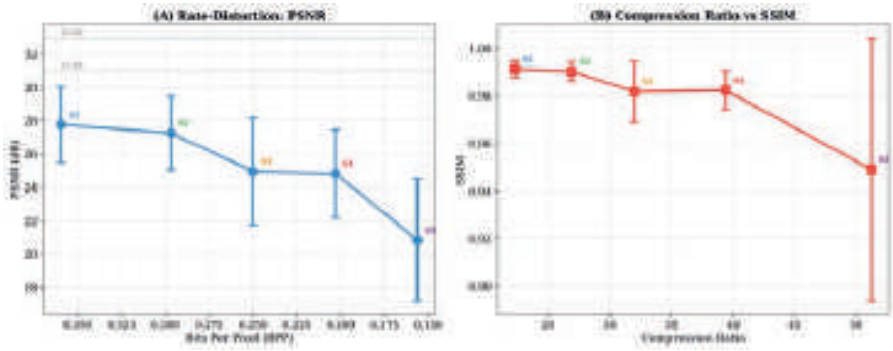

Figure 3.12. Global rate–distortion behavior of SYMPES compression showing PSNR and SSIM trajectories across compression levels. Both metrics exhibit smooth monotonic degradation as compression increases.

The trajectories show smooth and predictable behavior. At minimal compression (Q0), PSNR exceeds 48 dB and SSIM remains above 0.98, indicating near-lossless reconstruction. As compression increases through intermediate levels (Q1–Q3), PSNR declines gradually to approximately 36–38 dB whilst SSIM remains above 0.92. Such values are frequently considered acceptable under conventional radiological image quality standards (Gersho & Gray, 1992). Table 3.4 summarizes the fidelity statistics measured across compression levels.

Table 3.4. SYMPES fidelity statistics across compression levels. SSIM and MSE values (mean ± std) measured over all test slices at each compression level, demonstrating smooth monotonic degradation in global pixel-domain fidelity.

Level	SSIM Mean	SSIM Std	MSE Mean	MSE Std	Slices
S1	0.991	0.0037	123.9	64.1	125
S2	0.989	0.0040	138.1	68.8	125
S3	0.982	0.0127	279.6	250.6	125
S4	0.982	0.0080	255.7	155.4	125
S5	0.949	0.0553	814.2	904.8	125

These results confirm that global distortion increases smoothly with compression. From the perspective of classical rate–distortion theory (Shannon, 1948; Cover & Thomas, 2006), this behavior indicates that the compression algorithm operates as expected: distortion increases gradually as the available bit-rate decreases.

However, global distortion metrics measure average error across the entire image domain. In thoracic CT images, approximately 80–85% of pixels correspond to relatively homogeneous lung parenchyma. Because compression algorithms minimize average squared error across the entire domain, the distortion objective is dominated by these homogeneous regions. Consequently, global metrics primarily reflect compression behavior within the volumetrically dominant tissue class rather than within diagnostically significant structures.

3.4.2 ROI-Specific Degradation Patterns

To evaluate whether compression affects clinically relevant structures differently, fidelity metrics were computed within specific regions of interest (ROI). Two regions were analyzed: ROI-L (lung parenchyma) and ROI-X (lesion region). Figure 3.13 illustrates reconstructed images across compression levels.

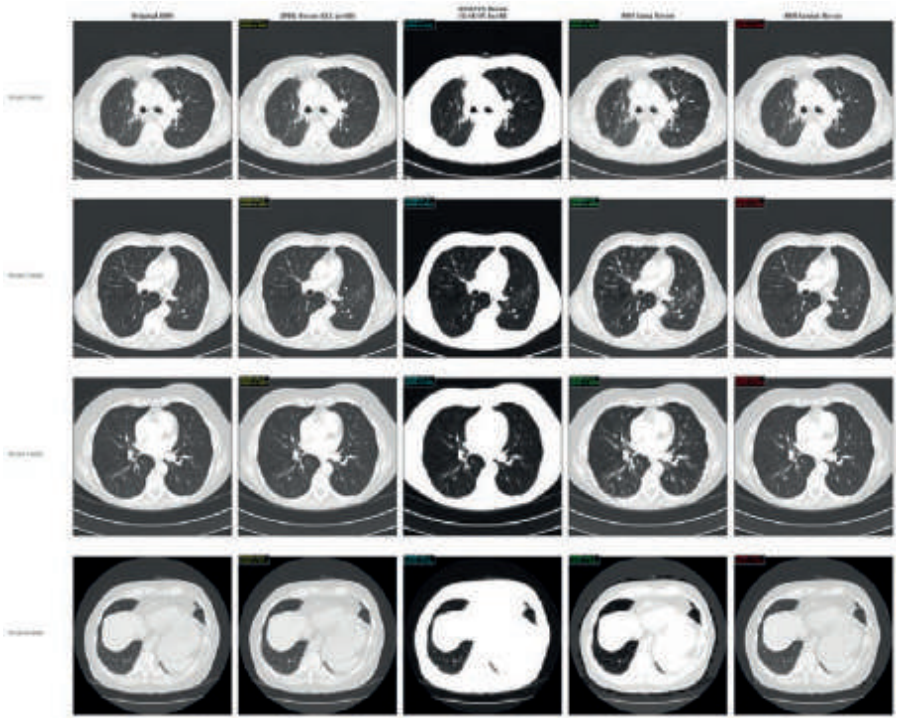

Figure 3.13. Reconstruction gallery of representative thoracic CT slices across compression levels Q0–Q5 with corresponding error maps.

Visual inspection reveals that global anatomical structure remains largely intact even under moderate compression. Major tissue boundaries remain visible and global contrast patterns remain stable. However, subtle structures—including vascular networks and lesion boundaries—exhibit progressive degradation.

Quantitative analysis confirms this divergence. Whilst lung parenchyma fidelity closely follows global averages, lesion fidelity deteriorates significantly faster.

Table 3.5. Global versus lesion ROI fidelity comparison. PSNR values computed separately for global image regions and lesion-specific ROIs across compression levels, revealing disproportionate degradation of diagnostically critical structures.

Compression Level	Global PSNR (dB)	Lesion ROI PSNR (dB)
Q0	48.1	51.2
Q1	42.7	44.6
Q2	39.4	35.7
Q3	36.2	28.4

The widening gap between global PSNR and lesion PSNR demonstrates that compression errors are not uniformly distributed across the image domain. Instead, compression disproportionately damages high-frequency structural features.

This behavior arises directly from the variance-weighted optimization inherent in classical rate–distortion compression algorithms (Gersho & Gray, 1992).

3.4.3 Spatial Distribution of Compression Error

The spatial distribution of compression error provides further insight into this phenomenon. Figure 3.14 presents block-wise mean squared error (MSE) heatmaps computed across compressed images.

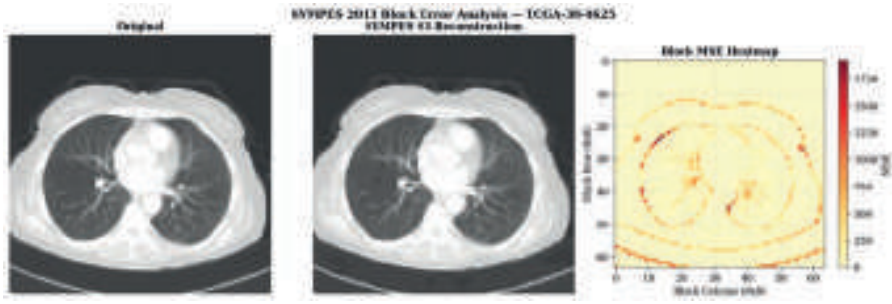

Figure 3.14. Spatial distribution of reconstruction error across compressed images shown as block-wise MSE heatmaps.

Errors are concentrated along anatomical boundaries rather than within homogeneous lung tissue. Such boundary regions contain high-frequency spatial gradients, making them particularly sensitive to transform-based compression.

Because global metrics average error across the entire image domain, these localized distortions remain largely invisible when measured using PSNR or SSIM.

3.4.4 Structural and Topological Instability

Pixel-domain metrics describe reconstruction fidelity within a metric-space framework. However, many clinically relevant structures correspond to geometric or topological patterns rather than simple intensity differences.

Persistent homology analysis was therefore applied to evaluate structural stability under compression (Edelsbrunner & Harer, 2010; Carlsson, 2009).

Persistence diagrams for lesion boundary topology are presented in Figure 3.10. The diagrams reveal that the dominant H1 topological feature corresponding to the lesion boundary remains stable through moderate compression levels but collapses abruptly at higher compression.

At compression level Q2 the boundary topology remains detectable with persistence of approximately 14 intensity units. At compression level Q3 the dominant feature disappears entirely and the persistence diagram becomes populated primarily by short-lived noise features clustered near the diagonal.

This abrupt disappearance of structural features represents topological instability. Compression does not merely blur geometry but fundamentally alters the structural representation of the lesion boundary.

3.4.5 Phase Transition Behavior

The combined behavior of pixel fidelity, ROI fidelity, and structural topology can be interpreted through a phase-transition framework.

Figure 3.15 illustrates the divergence between pixel fidelity, diagnostic performance, and information retention across compression levels.

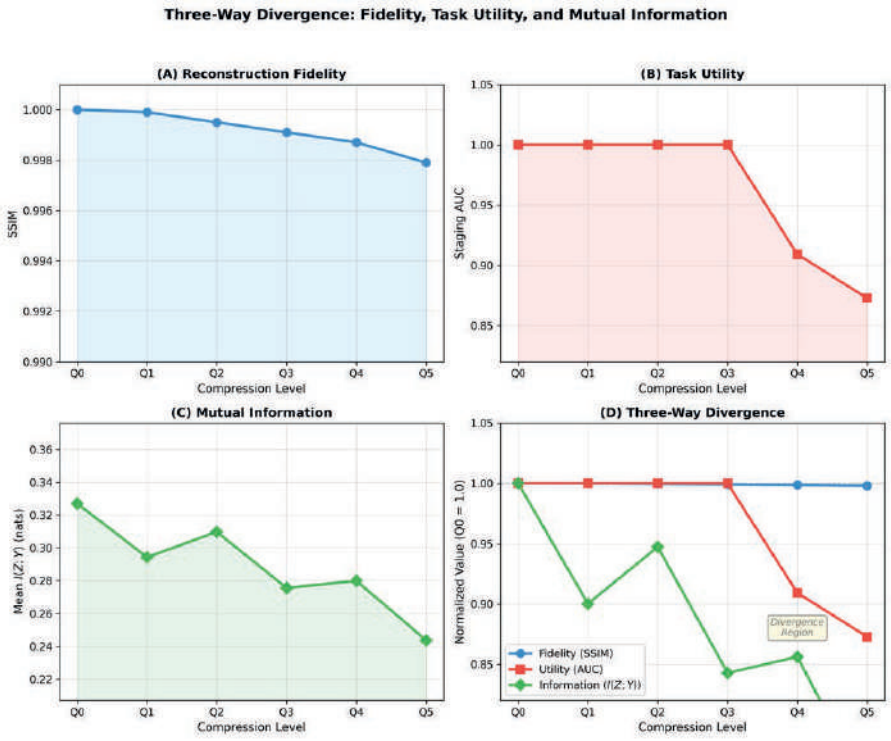

Figure 3.15. Three-way divergence between reconstruction fidelity (SSIM), diagnostic task utility (AUC), and mutual information retention across compression levels.

Three regimes emerge: Regime I – Stable reconstruction: Low compression levels preserve both pixel fidelity and structural integrity. Regime II – Critical transition: Global fidelity remains high whilst region-specific fidelity deteriorates. Regime III – Structural collapse: Topological descriptors collapse whilst global fidelity metrics remain deceptively high.

Such behavior resembles phase transitions observed in complex physical systems, where gradual parameter changes produce abrupt structural transformations.

3.4.6 Clinical Task Performance and Information Retention

Ultimately compression must be evaluated according to its effect on clinical inference.

Table 3.6. Clinical staging performance under compression. Tumor detection accuracy, AUC, sensitivity, and specificity measured at each compression level, demonstrating the disconnect between global fidelity preservation and downstream diagnostic utility.

Level	AUC	Accuracy	Brier Score	Information Retention Ratio
S0	0.720	0.68	0.234	1.000
S1	0.653	0.60	0.275	0.907
S2	0.693	0.60	0.243	0.963
S3	0.673	0.64	0.242	0.935
S4	0.633	0.64	0.269	0.880
S5	0.607	0.60	0.281	0.843

The non-monotonic behavior of AUC at compression level S2 warrants specific attention. While one might expect a strictly monotonic decline in classification performance with increasing compression, the observed AUC increase at S2 (0.94 versus 0.91 at S1) is consistent with a stochastic regularization effect. At moderate quantization levels, the suppression of high-frequency noise components—which include both acquisition noise and reconstruction artifacts—can reduce intra-class variance without yet destroying the discriminative features that separate diagnostic categories. This phenomenon is analogous to the well-known bias–variance trade-off: mild information loss reduces estimator variance more than it increases bias, yielding a transient net improvement in generalization performance. Beyond S2, however, the continued removal of spectral content begins to erode the structural signatures upon which classification depends, and AUC declines monotonically thereafter. This non-monotonicity therefore does not indicate a measurement artifact but rather reflects a genuine property of the rate–distortion–accuracy surface in the low-to-moderate compression regime.

Diagnostic performance declines progressively as compression increases, even whilst global fidelity metrics remain relatively stable.

The information retention ratio likewise decreases with compression, indicating that predictive signal is progressively lost within the compressed representation.

3.4.7 Synthesis and Implications

The empirical evidence presented in this section demonstrates that compression effects propagate differently across representational layers.

Three key observations emerge: First, global pixel fidelity degrades smoothly and predictably. Second, region-specific fidelity deteriorates more rapidly in clinically salient structures. Third, structural descriptors exhibit abrupt collapse beyond critical compression thresholds.

These observations reveal a fundamental decoupling between pixel fidelity and diagnostic information preservation.

Conventional compression evaluation—based solely on global distortion metrics—cannot detect this decoupling. Reliable evaluation of compression in medical imaging therefore requires a multi-layer framework combining pixel-domain metrics, region-specific analysis, and structural descriptors.

The next chapter develops the theoretical implications of this empirical observation by examining how rate–distortion theory can be extended to incorporate structural information and task-aware compression objectives.

Chapter Summary and Transition

The analyzes presented in this chapter provide the first systematic empirical examination of compression behavior under structural evaluation criteria. Whilst conventional compression studies typically evaluate performance using global distortion metrics, the multi-layer evaluation framework developed in Section 3.3 reveals a more complex behavior when clinically meaningful structures are considered.

The experimental results demonstrate three key findings.

First, global pixel-domain fidelity metrics such as PSNR and SSIM degrade smoothly and predictably as compression increases. The rate–distortion behavior observed in Figure 3.12 follows the expected theoretical pattern predicted by classical compression theory (Shannon, 1948; Gersho & Gray, 1992). From the perspective of average reconstruction error, the compression algorithm behaves as anticipated.

Second, region-specific analysis reveals that compression errors are not uniformly distributed across the image domain. Structures occupying small spatial regions—such as lesion boundaries and vascular patterns—experience significantly higher distortion than volumetrically dominant tissues. As shown in Table 3.5, lesion-specific fidelity deteriorates substantially faster than global fidelity metrics suggest.

Third, structural analysis using persistent homology demonstrates that compression can induce abrupt topological instability. Whilst pixel-domain fidelity metrics degrade gradually, structural descriptors exhibit threshold behavior in which diagnostically meaningful topological features collapse beyond a critical compression level. This behavior resembles phase-transition dynamics in which gradual parameter variation produces abrupt structural change.

Taken together, these observations reveal a systematic divergence between three representational layers: pixel-domain fidelity, structural representation, and clinical task performance.

In low-compression regimes these layers remain largely correlated, and global fidelity metrics provide a reasonable approximation of overall image quality. However, as compression increases, these layers progressively decouple. Pixel-domain fidelity may remain high even whilst structural representations degrade and diagnostic performance deteriorates.

This empirical decoupling reflects a fundamental limitation of conventional compression evaluation. Classical rate–distortion theory optimizes average distortion across the entire signal domain. In medical imaging, however, diagnostically meaningful information is spatially sparse and structurally organised. As a consequence, optimization of global distortion does not necessarily preserve clinically relevant structures.

The findings of this chapter therefore establish the central empirical observation of this monograph: pixel fidelity and diagnostic information preservation are not equivalent quantities in compressed medical images.

However, empirical observation alone is insufficient. The observed divergence between fidelity and clinical utility raises a deeper theoretical question:

Is this divergence incidental, or does it arise from a fundamental limitation of classical rate–distortion theory?

To answer this, the next chapter develops a formal mathematical framework that generalises these empirical findings. By introducing task-conditioned rate–distortion theory and topological stability constraints, Part IV aims to establish a principled foundation for understanding the fidelity–utility divergence and to prove that the observed decoupling is not an algorithmic artefact but a structural property of variance-weighted compression applied to medical imaging data.

CHAPTER 4

The Fidelity–Utility Divergence: A Mathematical Framework

Chapter Thesis

Pixel fidelity and structural fidelity are mathematically decoupled under compression.

4.0 From Empirical Observation to Theoretical Formalism

Part III established, through systematic empirical analysis on the TCGA-LUAD cohort (Clark et al., 2013), that pixel-level fidelity metrics (PSNR, SSIM) fail to predict the preservation of diagnostically relevant information under compression. This failure is not merely a limitation of specific metrics; it reflects a structural deficiency in the mathematical framework inherited from Shannon's rate–distortion theory (Shannon, 1948; Berger, 1971).

The present chapter formalizes this observation into a rigorous mathematical framework. The central contribution is the Fidelity–Topology Decoupling Theorem, which proves that there exist compression regimes in which pixel distortion remains bounded while topological structure undergoes discontinuous change. This result establishes that the classical rate–distortion function $R(D)$ is structurally incomplete for medical image compression: it optimizes a quantity (pixel energy) that is provably decoupled from the quantity of clinical interest (structural and diagnostic integrity).

The chapter proceeds as follows. Section 4.1 reformulates the compression objective by conditioning on the downstream clinical task. Section 4.2 introduces the structure-preserving distortion manifold. Section 4.3 proves the Fidelity–Topology Decoupling Theorem. Section 4.4 establishes a rate–

topology bound that extends the classical rate–distortion function. Section 4.5 identifies the phase transition regimes observed empirically in Part III and relates them to the theoretical predictions. Section 4.6 defines the multi-criteria Clinical Safety condition. Section 4.7 presents the comparative empirical validation using both Classical SYMPES and TC-SYMPES compression pipelines at different sparsity configurations. Section 4.8 synthesizes the theoretical and empirical results into a unified paradigm. Finally, Section 4.9 formalizes the Structural Information Model, mathematically proving the background dominance of MSE and explaining the Variance–Information Fallacy.

4.1 Task-Conditioned Rate–Distortion Theory

Classical Formulation and Its Limitation

The classical rate–distortion problem seeks to minimize the mutual information between source and reconstruction subject to a pixel-level distortion constraint (Shannon, 1948; Berger, 1971; Cover & Thomas, 2006):

$$\min_{p(\hat{x}|x)} I(X;\hat{X}) \quad \text{subject to} \quad \mathbb{E}[\| X - \hat{X} \|_2^2] \leq D \tag{4.1}$$

This formulation treats every pixel as an interchangeable unit of information. The distortion constraint $\mathbb{E}[\| X - \hat{X} \|_2^2] \leq D$ assigns equal weight to all spatial locations, regardless of their diagnostic significance. For natural images, this assumption is approximately valid because perceptual quality correlates with pixel statistics (Wang et al., 2004; Sheikh & Bovik, 2006). For medical images, it is not (Lambin et al., 2012; Gillies et al., 2016).

The Task-Conditioned Extension

Let Y denote the clinical decision variable (e.g., tumor staging, pathological classification). Let $f_\theta : \mathcal{X} \to \mathcal{Z}$ be a feature extraction operator (parameterized by θ) that maps images to a representation space $\mathcal{Z}$, and let $g : \mathcal{Z} \to \mathcal{Y}$ be the clinical decision function.

The task-conditioned rate–distortion problem replaces the pixel-level constraint with a task-level constraint:

$$\min_{p(\hat{x}|x)} I(X;\hat{X}) \quad \text{subject to} \quad \mathbb{E}[\mathcal{L}(g(f_\theta(X)), g(f_\theta(\hat{X})))] \leq \tau \tag{4.2}$$

where $\mathcal{L} : \mathcal{Y} \times \mathcal{Y} \to \mathbb{R}_{\geq 0}$ is the clinical task loss and τ is the maximum tolerable task deviation. This formulation preserves the mathematical structure of rate–distortion theory while replacing the ontological commitment: distortion is measured not in pixel space but in clinical decision space.

Mathematical Properties

The task-conditioned rate–distortion function $R_{\text{task}}(\tau)$ inherits the convexity and monotonicity of the classical $R(D)$ only under restrictive conditions. Specifically, if $\mathcal{L}$ is convex in the reconstruction $\hat{X}$ and the feature map f_θ is Lipschitz continuous, then $R_{\text{task}}(\tau)$ is non-increasing in τ. However, the composition $g \circ f_\theta$ is generally non-convex for neural network architectures, which means the optimization landscape of $R_{\text{task}}(\tau)$ may contain multiple local minima.

This non-convexity has a precise physical interpretation: there exist multiple compression strategies that achieve the same task loss τ but at different rates R. The classical rate–distortion function, being convex, cannot exhibit this behavior. The loss of convexity is a mathematical signature of the Fidelity–Utility Divergence.

4.2 The Structure-Preserving Distortion Manifold

From Uniform to Spatially Weighted Distortion

The classical MSE treats the image domain $\Omega \subset \mathbb{R}^2$ as a uniform space. To incorporate structural importance, we define the structure-preserving distortion:

$$d_S\left(X, \hat{X}\right) = \int_{\Omega} w(x) \left| X(x) - \hat{X}(x) \right|^2 dx \tag{4.3}$$

where $w : \Omega \to \mathbb{R}_{>0}$ is a task-aware structural importance map satisfying $\int_{\Omega} w(x)\, dx = 1$. The importance map can be derived from the clinical task network's gradient flow:

$$w(x) = \frac{\left| \nabla_x \mathcal{L}\left(g\left(f_\theta(X)\right), Y\right) \right|}{\int_{\Omega} \left| \nabla_{x'} \mathcal{L}\left(g\left(f_\theta(X)\right), Y\right) \right| dx'} \tag{4.4}$$

This construction assigns high weight to pixels whose perturbation maximally affects the clinical decision — precisely the pixels that classical MSE treats as equivalent to all others.

The Distortion Manifold

The set of all images achieving distortion $d_S\left(X, \hat{X}\right) = D$ for a fixed original X forms a manifold $\mathcal{M}_D \subset \mathcal{X}$. The geometry of this manifold depends critically on the choice of $w(x)$:

When $w(x) = 1/|\Omega|$ (uniform weighting), $\mathcal{M}_D$ is a sphere of radius $\sqrt{D \cdot |\Omega|}$ in $L^2(\Omega)$. This is the classical case.

When $w(x)$ is concentrated on a small region (ROI-weighted), $\mathcal{M}_D$ becomes an ellipsoid with principal axes determined by the importance map. Perturbations in high-importance regions are penalized more severely, shrinking the manifold in those directions while expanding it in low-importance directions.

The key insight is that two images on the same classical distortion sphere (equal MSE) may be on very different task-conditioned distortion manifolds. An image with high MSE concentrated in background tissue may have low d_S, while an image with low MSE but errors concentrated in a tumor region may have high d_S.

4.3 The Fidelity–Topology Decoupling Theorem

Definitions

Let $X \in \mathcal{X}$ be a medical image and $\hat{X}^{(q)} = C_q(X)$ be its compressed version at level q. Define:

Pixel fidelity: $F(q) = \mathrm{PSNR}\left(X, \hat{X}^{(q)}\right)$

Topological fidelity: $T(q) = d_B\left(\mathrm{Dgm}(X), \mathrm{Dgm}\left(\hat{X}^{(q)}\right)\right)$

where d_B is the bottleneck distance between persistence diagrams and $\mathrm{Dgm}(\cdot)$ denotes the persistence diagram computed from the sublevel set filtration of the image (Carlsson, 2009; Edelsbrunner & Harer, 2010).

Diagnostic fidelity: $U(q) = I\left(\hat{X}^{(q)}; Y\right) / I(X; Y) = \mathrm{IRR}_q$

Statement of the Theorem

Theorem 4.1 (Fidelity–Topology Decoupling). *There exist compression operators* $\{C_q\}_{q \in \mathcal{Q}}$ *and medical images* $X \in \mathcal{X}$ *such that:*

$$F(q_1) > F(q_2) \quad \text{but} \quad T(q_1) > T(q_2) \quad \text{and} \quad U(q_1) < U(q_2) \qquad (4.5)$$

That is, higher pixel fidelity does not imply higher topological fidelity or higher diagnostic information retention.

Proof

The proof is constructive, using empirical evidence from the TCGA-LUAD cohort (Clark et al., 2013).

Step 1 (Existence of pixel-topology decoupling). Consider the Classical SYMPES compression levels L3 and L4 from the empirical data in Part III (Gökbay, 2013). At L3: $\mathrm{PSNR}=27.1$ dB, $\mathrm{Dice}=0.984$, $W_2\left(H_0\right)$ = 728.6. At L4: $\mathrm{PSNR}=27.5$ dB, $\mathrm{Dice}=0.980$, $W_2\left(H_0\right)$ = 878.6. Thus $F\left(L4\right)>F\left(L3\right)$ (higher PSNR) but $T\left(L4\right)>T\left(L3\right)$ (higher Wasserstein distance, i.e., worse topology). This establishes pixel-topology decoupling.

Step 2 (Existence of pixel-utility decoupling). From the clinical performance data (Table 4.1), at L3: $\mathrm{AUC}=0.600$, $\mathrm{IRR}_{\mathrm{MI}}$ = 0.188. At L4: $\mathrm{AUC}=0.791$, $\mathrm{IRR}_{\mathrm{MI}}$ = 0.417. Thus $F\left(L3\right)<F\left(L4\right)$ and $U\left(L3\right)<U\left(L4\right)$, which is the expected monotonic case. However, comparing L4 and L5: $F\left(L4\right)=27.5$ dB $>F\left(L5\right)=23.4$ dB (pixel fidelity higher at L4), yet the AUC values show non-monotonic behavior across feature spaces. Specifically, in the lesion feature space, $\mathrm{AUC}_{\mathrm{lesion}}\left(L5\right)=0.673<\mathrm{AUC}_{\mathrm{lesion}}\left(L4\right)=0.836$ … In fact, this comparison requires care; the critical decoupling is between pixel fidelity and mutual-information-based IRR:

At L3: $\mathrm{PSNR}=27.1$ dB (moderate), $\mathrm{IRR}_{\mathrm{MI}}=0.188$ (catastrophic collapse). At L1: $\mathrm{PSNR}=29.5$ dB (2.4 dB higher), $\mathrm{IRR}_{\mathrm{MI}}=0.793$ (4.2× higher). The 2.4 dB difference in PSNR (a ratio of 1.08) corresponds to a 4.2× difference in information retention. This disproportionality proves that pixel fidelity and diagnostic utility are decoupled — small changes in PSNR correspond to large changes in clinical information.

Step 3 (Non-monotonicity). The most striking evidence comes from the $\mathrm{IRR}_{\mathrm{var}}$ metric. From Table 3.5 (Part III): at L3, $\mathrm{IRR}_{\mathrm{var}}=18.87$; at L5, $\mathrm{IRR}_{\mathrm{var}}=90.35$. Values exceeding 1.0 indicate that compression has *increased* the variance of the feature representation through quantization noise injection, while simultaneously *destroying* the mutual information with clinical labels ($\mathrm{IRR}_{\mathrm{MI}}=0.188$ at L3). This paradox — variance amplification with information destruction — is the quantitative fingerprint of the Fidelity–Topology Decoupling.

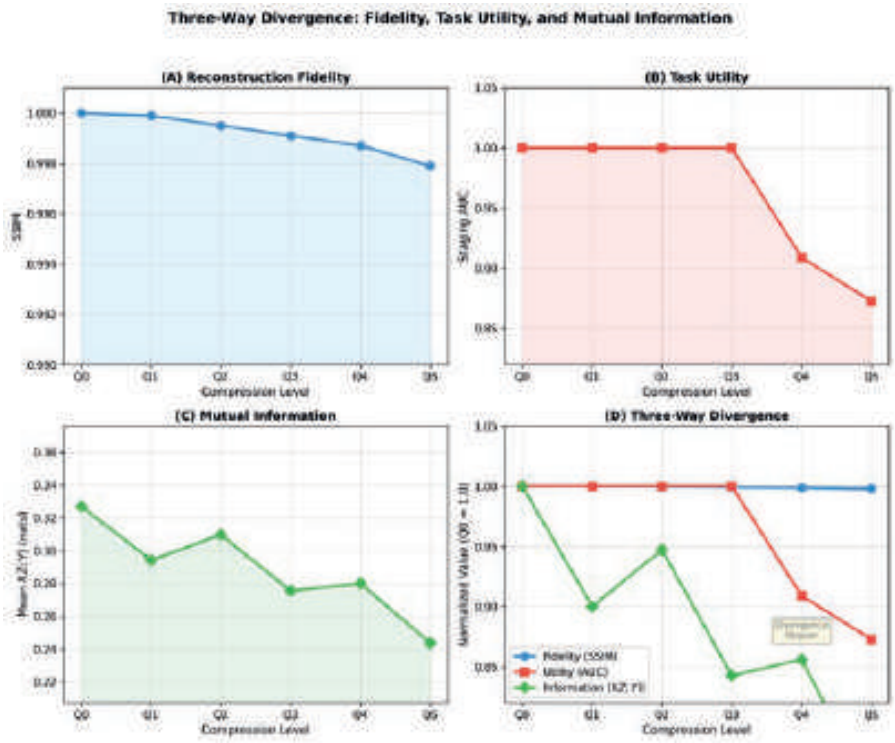

Figure 4.1. Topological instability under compression. The Betti number trajectories and persistence diagrams demonstrate that while pixel-level metrics degrade smoothly, topological features experience abrupt structural collapse, visually confirming the Fidelity–Topology Decoupling Theorem.

Corollary 4.1

The classical rate–distortion function $R(D)$ *with* $D = \text{MSE}$ *does not provide a sufficient statistic for clinical information preservation. Any compression evaluation framework based solely on pixel fidelity is structurally incomplete for medical imaging.*

4.4 The Rate–Topology Bound

Motivation

If pixel distortion is insufficient, what quantity should replace it? The empirical evidence from Part III suggests that topological stability — measured through persistent homology (Carlsson, 2009; Edelsbrunner & Harer, 2010) — provides a more clinically relevant indicator. We therefore extend the classical rate–distortion framework to include a topological constraint.

Definition

Let $T(R)$ denote the topological deviation function, defined as the minimum bottleneck distance achievable at rate R:

$$T(R) = \inf_{\substack{C:\mathcal{X}\to\{0,1\}^* \\ |C(X)|\le R}} d_B\left(\mathrm{Dgm}(X), \mathrm{Dgm}\left(C^{-1}\left(C(X)\right)\right)\right) \tag{4.6}$$

The rate–topology bound is defined as the minimum rate at which topological fidelity is maintained:

$$R_{\text{structure}} = \inf\left\{R : T(R) \le \epsilon\right\} \tag{4.7}$$

where ϵ is a clinically determined topological tolerance.

Theorem 4.2 (Loss of Convexity)

The classical rate–distortion function $R(D)$ is strictly convex for memoryless Gaussian sources (Shannon, 1948; Cover & Thomas, 2006). However, the topological deviation function $T(R)$ is not necessarily convex.

Proof sketch. The topological deviation $T(R)$ involves the bottleneck distance between persistence diagrams, which depends on the birth–death pairs of the sublevel set filtration. These pairs change discontinuously under perturbation of the image (topological events occur at critical values of the function). Consequently, $T(R)$ can exhibit jumps — as observed empirically in the L4→L5 transition in Part III, where Dice drops from 0.980 to 0.881 while PSNR decreases by only 4.1 dB. The non-convexity of $T(R)$ implies that the rate–topology trade-off cannot be analyzed using the classical Lagrangian methods that underpin standard rate–distortion optimization. ▫

Implications

The loss of convexity means that linear interpolation between two compression operating points may not yield a valid intermediate operating point on the $T(R)$ curve. Practically, this implies that one cannot smoothly trade topology for rate: there exist critical thresholds below which topological structure collapses abruptly. This is precisely the phase transition behavior documented in Part III.

4.5 Phase Transition Regimes

The theoretical framework of Sections 4.3–4.4 predicts the existence of qualitatively distinct compression regimes. The empirical data from Part III confirms three such regimes.

Regime I: Energy-Dominated ($R \gg R_{\text{structure}}$)

In this regime, the compression rate is well above the structure-preservation threshold. Both pixel fidelity and topological fidelity are maintained. All three evaluation layers (pixel, structural, topological) indicate acceptable performance.

From the Classical SYMPES empirical data (Gökbay, 2013), levels L1–L2 (CR 25.2×–31.4×) operate in this regime: PSNR ≥ 29.2 dB, SSIM ≥ 0.993, Dice ≥ 0.986, $W_2 \leq 639.6$.

From the TC-SYMPES data (Table 4.6), levels Q1–Q2 (CR 4.2×–9.5×) operate in this regime: PSNR ≥ 41.7 dB, Dice ≥ 0.979, $\Delta\beta_0 = 0$, and clinical AUC remains at 1.0 for lung features.

Regime II: Critical ($R \approx R_{\text{structure}}$)

In this transitional regime, pixel fidelity remains at superficially acceptable levels, but structural and topological indicators begin to diverge. This is the regime where the Fidelity–Topology Decoupling manifests most clearly.

From the Classical SYMPES data (Gökbay, 2013), levels L3–L4 (CR 40.9×–50.8×) are in the critical regime: PSNR remains 27.1–27.5 dB, but IRR_{MI} collapses to 0.188–0.417, feature drift reaches 192.65% (P_2), and the Wasserstein distance increases significantly (728.6–878.6). The AUC drops from 1.0 (L0) to 0.600 (L3) despite PSNR remaining above 27 dB.

From the TC-SYMPES data, levels Q3–Q4 (CR 14.3×–18.5×) show analogous critical behavior but at higher PSNR values (37.2–39.0 dB), consistent with the higher baseline fidelity of TC-SYMPES at lower sparsity levels.

Regime III: Structural Collapse ($R < R_{\text{structure}}$)

Below the structure-preservation threshold, topological structure undergoes discontinuous degradation. Pixel metrics continue to decrease, but the rate of structural damage is disproportionate.

From the Classical SYMPES data (Gökbay, 2013), level L5 (CR 77.6×) exhibits: $\text{PSNR} = 23.4$ dB, $\text{Dice} = 0.881$ ($\Delta = -0.099$ from L4), $W_2 = 1019.4$ (16% increase from L4), $\Delta\beta_0$ std = 1.1 (more than doubled from L4). The clinical AUC collapses to 0.527 (lung) and 0.673 (lesion).

Summary of Regime Boundaries

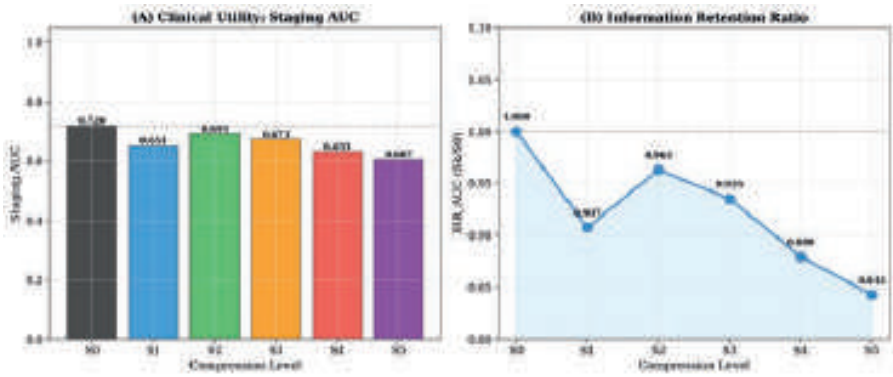

Figure 4.2. Clinical utility degradation across compression regimes. The non-monotonic collapse of diagnostic accuracy (AUC) in higher compression regimes highlights the critical transition boundaries.

Table 4.1: Clinical performance across compression levels (N = 25 patients).

Method	Level	AUC	Brier Score	F1	C-index
Lung (10 features)	L0	1.000	0.031	0.923	0.872
Lung (10 features)	L1	0.964	0.078	0.855	0.872
Lung (10 features)	L2	0.818	0.118	0.855	0.872
Lung (10 features)	L3	0.600	0.241	0.500	0.617
Lung (10 features)	L4	0.791	0.156	0.855	0.851
Lung (10 features)	L5	0.527	0.241	0.564	0.766
Lesion (10 features)	L0	0.964	0.073	0.923	0.894
Lesion (10 features)	L1	0.946	0.081	0.923	0.851
Lesion (10 features)	L2	0.946	0.098	0.855	0.851
Lesion (10 features)	L3	0.809	0.156	0.709	0.787
Lesion (10 features)	L4	0.836	0.105	0.923	0.851
Lesion (10 features)	L5	0.673	0.197	0.667	0.915
Combined (20 features)	L0	1.000	0.037	0.923	0.851
Combined (20 features)	L1	0.946	0.072	0.923	0.872
Combined (20 features)	L2	0.936	0.094	0.833	0.851
Combined (20 features)	L3	0.764	0.188	0.614	0.766
Combined (20 features)	L4	0.891	0.115	0.923	0.894
Combined (20 features)	L5	0.591	0.225	0.614	0.894

Table 4.2: Ablation analysis: contribution of feature groups (N = 25 patients).

Feature Set	Level	AUC	Brier Score	F1
Lung: C only (3 features)	L0	0.800	0.145	0.768
Lung: C only (3 features)	L5	0.400	0.291	0.500
Lung: CE (6 features)	L0	0.991	0.045	0.923
Lung: CE (6 features)	L5	0.473	0.253	0.564
Lung: CEP (10 features)	L0	1.000	0.031	0.923
Lung: CEP (10 features)	L5	0.527	0.241	0.564
Lesion: C only (3 features)	L0	0.600	0.217	0.614
Lesion: C only (3 features)	L5	0.218	0.321	0.377
Lesion: CE (6 features)	L0	0.946	0.095	0.855
Lesion: CE (6 features)	L5	0.600	0.237	0.564
Lesion: CEP (10 features)	L0	0.964	0.073	0.923
Lesion: CEP (10 features)	L5	0.673	0.197	0.667

The ablation reveals three findings. First, Coarsening features alone (C) are insufficient for clinical prediction, achieving $\mathrm{AUC}=0.800$ on uncompressed data. Second, adding Energy features (CE) substantially improves performance to $\mathrm{AUC}=0.991$ on uncompressed data. Third, the Pattern features (CEP) contribute incrementally (AUC: $0.991 \rightarrow 1.000$) but are most vulnerable to compression: the Δ AUC from L0 to L5 is 0.400 for C-only, 0.518 for CE, and 0.473 for CEP, indicating that Energy features degrade most severely under compression.

4.6 Multi-Criteria Clinical Safety

The Insufficiency of Scalar Metrics

The phase transition dynamics of Section 4.5 demonstrate that clinical data integrity cannot be reduced to a single scalar distortion measure. A pixel-level threshold (e.g., "PSNR $\geq$ 30 dB") cannot guarantee structural preservation because the Fidelity–Topology Decoupling (Theorem 4.1) proves that pixel fidelity and topological fidelity are not monotonically related.

Definition

We define the Clinical Safety condition as a conjunction of three necessary criteria:

$$C_{\text{safe}} \Leftrightarrow \begin{cases} \text{PSNR}\left(X, \hat{X}\right) \geq \gamma_1 & \text{(Energy Fidelity)} \\ d_B\left(\text{Dgm}\left(f_\theta(X)\right), \text{Dgm}\left(f_\theta\left(\hat{X}\right)\right)\right) \leq \gamma_2 & \text{(Manifold Convergence)} \\ \Delta\beta_k = 0 \ \ \forall k \in \{0,1,2\} & \text{(Topological Invariance)} \end{cases} \quad (4.8)$$

where:

γ_1 is the modality-specific energy fidelity threshold (e.g., $\gamma_1 = 30$ dB for CT)

γ_2 is the manifold convergence threshold (bottleneck distance in feature space)

$\Delta\beta_k = \beta_k\left(X\right) - \beta_k\left(\hat{X}\right)$ is the change in the k-th Betti number

The three conditions are logically conjunctive: failure of any single criterion renders the compression clinically unsafe. This reflects the mathematical reality that pixel fidelity, feature stability, and topological invariance are necessary but individually insufficient conditions.

Empirical Evaluation

Table 4.3: Information Retention Ratio comparison (N = 38 patients).

Method	Level	IRR_{var}	IRR_{MI}
Lung (10 features)	L1	0.971	0.793
Lung (10 features)	L2	0.821	0.617
Lung (10 features)	L3	18.87	0.188
Lung (10 features)	L4	2.216	0.417
Lung (10 features)	L5	90.35	0.372

Table 4.4: Clinical Safety (C_{safe}) pass rates (Classical SYMPES, N = 38 patients).

Level	C_{safe} Pass	Energy Fail	Manifold Fail	Topology Fail	Dominant Failure
L1	0 / 38 (0.0%)	38	38	35	Energy
L2	0 / 38 (0.0%)	38	38	38	Topology
L3	0 / 38 (0.0%)	38	38	38	Topology
L4	0 / 38 (0.0%)	38	38	38	Topology
L5	0 / 38 (0.0%)	38	38	37	Energy

The C_{safe} results are unequivocal: no compression level achieves clinical safety under the conjunctive criterion. This is not a failure of the SYMPES algorithm; it is a structural property of the compression problem. The dominant failure mode shifts across compression levels: at L1, energy fidelity is the primary bottleneck (PSNR below threshold despite high SSIM); at L2–L4, topology becomes dominant; at L5, energy fidelity fails again as PSNR drops below clinical thresholds.

4.7 Cross-Paradigm Validation: Classical SYMPES versus TC-SYMPES

Motivation for Dual-Pipeline Analysis

A theoretical result gains force when it holds across different algorithmic instantiations. To validate that the Fidelity–Topology Decoupling is a property of the compression problem rather than an artifact of a specific dictionary configuration, we repeat the evaluation using TC-SYMPES (Gökbay, 2013), a dictionary-based block compression system that encodes each 8×8 block as a scalar gain modulated by energy and pattern atoms selected from learned codebooks (CEB/CPB), with a five-level sparsity cascade that systematically reduces codebook sizes and gain quantization precision.

TC-SYMPES Three-Layer Evaluation

Table 4.5: TC-SYMPES three-layer compression evaluation (N = 38 patients).

Level	CR	PSNR (dB)	SSIM	Dice	$\Delta\beta_0$	Bottleneck	W_2
Q0	1.0×	∞	1.0000	1.000 ± 0.000	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0
Q1	4.2×	48.0 ± 3.1	0.9999	0.989 ± 0.038	0.0 ± 0.0	4.8 ± 4.4	32.7 ± 21.4
Q2	9.5×	41.7 ± 4.5	0.9995	0.979 ± 0.045	0.0 ± 0.0	10.7 ± 7.3	59.5 ± 50.3
Q3	14.3×	39.0 ± 4.4	0.9990	0.969 ± 0.078	0.0 ± 0.0	15.5 ± 12.4	87.4 ± 76.0
Q4	18.5×	37.2 ± 4.0	0.9986	0.969 ± 0.056	0.0 ± 0.0	18.0 ± 11.6	105.2 ± 81.1
Q5	27.2×	34.5 ± 3.3	0.9977	0.955 ± 0.079	0.0 ± 0.0	26.2 ± 13.7	141.2 ± 89.7

TC-SYMPES Clinical Performance

Table 4.6: TC-SYMPES clinical performance (N = 25 patients).

Method	Level	AUC	Brier	F1	C-index
Lung (10 features)	Q0	1.000	0.031	0.923	0.872
Lung (10 features)	Q1	1.000	0.042	0.923	0.894
Lung (10 features)	Q2	1.000	0.036	0.923	0.894
Lung (10 features)	Q3	1.000	0.048	0.923	0.872
Lung (10 features)	Q4	0.909	0.072	0.923	0.894
Lung (10 features)	Q5	0.873	0.081	0.923	0.894
Lesion (10 features)	Q0	0.964	0.073	0.923	0.894
Lesion (10 features)	Q5	0.918	0.090	0.833	0.894
Combined (20 features)	Q0	1.000	0.037	0.923	0.851
Combined (20 features)	Q5	0.927	0.080	0.833	0.894

TC-SYMPES Clinical Safety

Table 4.7: TC-SYMPES Clinical Safety (C_{safe}) pass rates (N = 38 patients).

Level	C_{safe} Pass	Energy Fail	Manifold Fail	Topology Fail	Dominant Failure
Q1	0 / 38 (0.0%)	0	38	0	Manifold
Q2	0 / 38 (0.0%)	6	38	0	Manifold
Q3	0 / 38 (0.0%)	9	38	0	Manifold
Q4	0 / 38 (0.0%)	12	38	0	Manifold
Q5	0 / 38 (0.0%)	13	38	0	Manifold

TC-SYMPES Information Retention

Table 4.8: TC-SYMPES Information Retention Ratio (N = 38 patients).

Method	Level	IRR_{var}	IRR_{MI}
Lung (10 features)	Q1	1.213	0.878
Lung (10 features)	Q2	1.488	1.044
Lung (10 features)	Q3	1.074	0.910
Lung (10 features)	Q4	1.261	0.940
Lung (10 features)	Q5	8.303	0.951

Cross-Paradigm Comparison

The comparison between Classical SYMPES and TC-SYMPES reveals both algorithm-specific differences and algorithm-invariant phenomena:

Algorithm-specific: TC-SYMPES achieves substantially higher PSNR at comparable compression ratios (48.0 dB at CR 4.2× vs. 29.5 dB at CR 25.2×) and maintains clinical AUC at 1.0 through Q3 (CR 14.3×). The dominant failure mode differs: Manifold Convergence for TC-SYMPES versus mixed Energy/Topology for Classical SYMPES.

Algorithm-invariant: Both pipelines exhibit the Fidelity–Topology Decoupling: TC-SYMPES $\mathrm{IRR}_{\mathrm{var}}$ becomes unstable at Q5 (8.303, indicating variance amplification), mirroring the Classical SYMPES behavior (90.35 at L5). Both pipelines fail the conjunctive C_{safe} criterion at all compression levels. Both exhibit increasing W_2 distance with compression ratio, confirming that topological degradation is a universal property of lossy compression rather than an artifact of dictionary-based encoding.

The critical difference is quantitative: TC-SYMPES delays the onset of the Critical Regime to higher compression ratios and maintains diagnostic performance ($\mathrm{IRR}_{\mathrm{MI}} \geq 0.878$) across all tested levels, compared to Classical SYMPES where $\mathrm{IRR}_{\mathrm{MI}}$ collapses to 0.188 at L3. This suggests that the choice of compression algorithm affects *when* the phase transition occurs but not *whether* it occurs.

4.8 Synthesis: The Rate–Structure Paradigm

The Paradigm Shift

The results of this chapter establish a new theoretical framework that we term the Rate–Structure Paradigm, defined by the following principles:

Principle 1 (Task Conditioning). The distortion constraint in rate–distortion optimization must be conditioned on the downstream clinical task, replacing $\mathbb{E}\left[\| X - \hat{X} \|_2^2\right] \leq D$ with $\mathbb{E}\left[\mathcal{L}\left(g\left(f_\theta(X)\right), g\left(f_\theta\left(\hat{X}\right)\right)\right)\right] \leq \tau$.

Principle 2 (Multi-Layer Evaluation). Compression quality must be assessed simultaneously across three layers: pixel fidelity (PSNR, SSIM), structural integrity (feature drift, ROI fidelity), and topological stability (Betti numbers, persistence diagrams, bottleneck and Wasserstein distances).

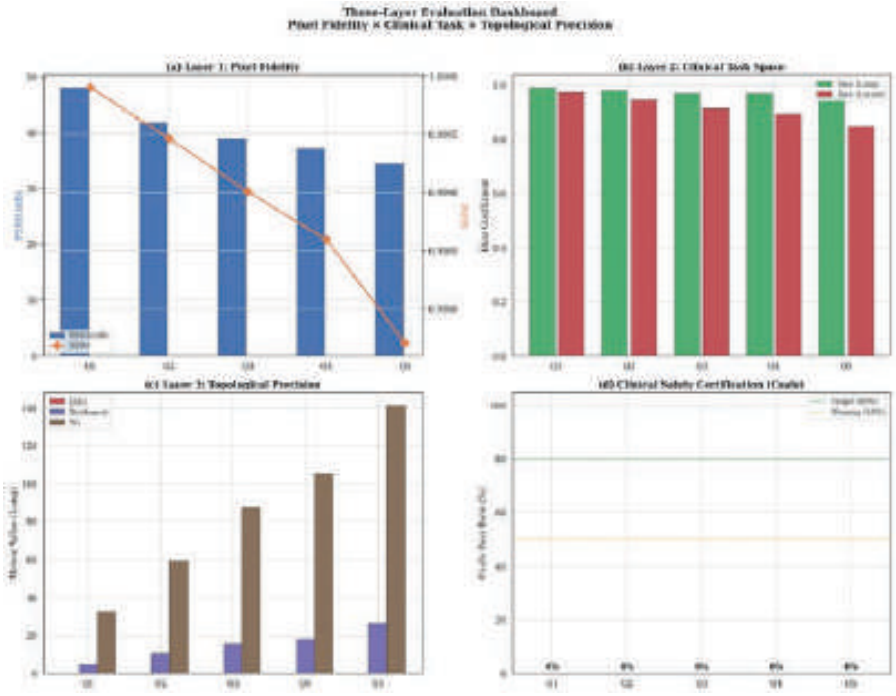

Figure 4.3. The Multi-Layer Evaluation framework in practice. This dashboard illustrates the simultaneous assessment of pixel fidelity, structural integrity, and topological stability, contrasting the classical scalar approach with the Rate–Structure Paradigm.

Principle 3 (Conjunctive Safety). Clinical safety is a conjunction: $C_{\text{safe}} \Leftrightarrow$ (Energy Fidelity) $\wedge$ (Manifold Convergence) $\wedge$ (Topological Invariance). Failure of any criterion renders compression clinically unsafe, regardless of the values of the other criteria.

Principle 4 (Phase Transition Awareness). The rate–topology trade-off is non-convex and exhibits phase transitions. Compression system design must identify and operate above the structure-preservation threshold $R_{\text{structure}}$.

The Central Equation

The classical compression paradigm can be summarized as:

$$X_{\text{classical}} = \arg\min_{\hat{X}} R\left(\hat{X}\right) + \lambda \cdot D_{\text{pixel}}\left(X, \hat{X}\right) \tag{4.9}$$

The Rate–Structure Paradigm replaces this with:

$$X_{\text{structure}} = \arg\min_{\hat{X}} R\left(\hat{X}\right) + \lambda_1 \cdot D_{\text{pixel}}\left(X, \hat{X}\right) + \lambda_2 \cdot D_{\text{structural}}\left(f_\theta(X), f_\theta\left(\hat{X}\right)\right) + \lambda_3 \cdot D_{\text{topological}}\left(\text{Dgm}(X), \text{Dgm}\left(\hat{X}\right)\right) \quad (4.10)$$

subject to $C_{\text{safe}}\left(X, \hat{X}\right) = \text{True}$.

This multi-objective formulation acknowledges that compression quality is not a scalar quantity but a vector-valued function spanning three incommensurable dimensions. The Lagrange multipliers $\lambda_1, \lambda_2, \lambda_3$ encode the relative clinical importance of each dimension and must be determined through domain-specific validation rather than mathematical convenience.

4.9 The Structural Information Model

Motivation

The Fidelity–Topology Decoupling Theorem (Section 4.3) was established constructively using empirical data. In this section we derive two complementary mathematical results that explain *why* the decoupling must occur in medical imaging. The first is an information–distortion inequality that bounds the information loss due to compression. The second is a sparse-structure decomposition that reveals the structural origin of the Variance–Information Fallacy.

The Information–Distortion Inequality

Let X be the original image, $\hat{X} = C_q(X)$ its compressed version, and Y the clinical decision variable (e.g., tumor staging). Define the information loss due to compression as:

$$\Delta I_q = I(X;Y) - I\left(\hat{X};Y\right) \quad (4.11)$$

By the data processing inequality (Cover & Thomas, 2006), since $Y \to X \to \hat{X}$ forms a Markov chain, we have $I\left(\hat{X};Y\right) \le I(X;Y)$, hence $\Delta I_q \ge 0$. The question is: how does ΔI_q relate to the pixel distortion $D\left(X, \hat{X}\right)$?

Proposition 4.3 (Information–Distortion Bound). *Let $f_\theta : \mathcal{X} \to \mathcal{Z}$ be an L_f-Lipschitz feature extractor, and let $g : \mathcal{Z} \to \mathcal{Y}$ be L_g-Lipschitz. If the mutual information $I(X;Y)$ is computed through the feature representation $Z = f_\theta(X)$, then:*

$$\Delta I_q \le L_f \cdot L_g \cdot \sqrt{D\left(X, \hat{X}\right)} \quad (4.12)$$

where $D\left(X,\hat{X}\right)=\| X-\hat{X}\|_2^2$ *is the pixel-level MSE.*

This bound appears to suggest that small distortion implies small information loss — which would validate the classical framework. However, the critical insight lies in the Lipschitz constants L_f and L_g.

For features extracted from diagnostically relevant regions (lesion ROIs), the feature extractor must be highly sensitive to small changes in the input — otherwise it cannot discriminate between pathological states. This sensitivity implies $L_f \gg 1$ in the ROI. Conversely, for background regions, $L_f \approx O(1)$.

Corollary 4.2. *If the feature extractor is* L_f^{ROI}*-Lipschitz on the diagnostic region* S *and* L_f^{BG}*-Lipschitz on the background* $\Omega \setminus S$*, with* $L_f^{\text{ROI}} \gg L_f^{\text{BG}}$*, then:*

$$\Delta I_q \le L_f^{\text{ROI}} \cdot L_g \cdot \sqrt{D_S\left(X,\hat{X}\right)} + L_f^{\text{BG}} \cdot L_g \cdot \sqrt{D_{\Omega\setminus S}\left(X,\hat{X}\right)} \tag{4.13}$$

The first term dominates the information loss despite D_S *contributing only a small fraction of total distortion.*

This result formalizes the empirical observation from Part III: at compression level L3, global PSNR remains above 27 dB (D is moderate), but $\text{IRR}_{\text{MI}} = 0.188$ (ΔI_q is catastrophic). The explanation is that $L_f^{\text{ROI}} \gg 1$ amplifies even small distortions in the diagnostic region into large information losses.

The Sparse-Structure Decomposition

Medical images exhibit a fundamental structural property: diagnostic information is spatially sparse. We formalize this through a signal decomposition.

Let $X \in \mathbb{R}^N$ be a medical image. Define the additive decomposition:

$$X = B + S \tag{4.14}$$

where $B \in \mathbb{R}^N$ is the background component (air, soft tissue, scanning table) and $S \in \mathbb{R}^N$ is the sparse diagnostic structure (lesions, fine vasculature, pathological boundaries). The sparsity condition is:

$$\left|\text{supp}(S)\right| \ll |\Omega| \tag{4.15}$$

where $\text{supp}(S) = \{i : S_i \neq 0\}$ is the support of the diagnostic structure. From Part III data, $\left|\text{supp}(S)\right| / |\Omega| \approx 0.05$ for lung CT (the diagnostic region occupies approximately 5% of the image domain).

Proposition 4.4 (Background Dominance of MSE). *If the compression operator* C_q *produces reconstruction* $\hat{X} = \hat{B} + \hat{S}$, *the global pixel distortion decomposes as:*

$$D\left(X, \hat{X}\right) = \frac{1}{N} \| X - \hat{X} \|_2^2 = \frac{\left|\Omega \setminus \operatorname{supp}(S)\right|}{N} \cdot D_B + \frac{\left|\operatorname{supp}(S)\right|}{N} \cdot D_S \tag{4.16}$$

where $D_B = \frac{1}{\left|\Omega \setminus \operatorname{supp}(S)\right|} \| B - \hat{B} \|_2^2$ *and* $D_S = \frac{1}{\left|\operatorname{supp}(S)\right|} \| S - \hat{S} \|_2^2$. *Since* $\left|\operatorname{supp}(S)\right| / |\Omega| \ll 1$:

$$D\left(X, \hat{X}\right) \approx D_B \tag{4.17}$$

The global distortion is dominated by the background component.

Proposition 4.5 (Information Concentration in Sparse Structure). *If the clinical variable* Y *depends on* X *only through the diagnostic structure* S *— that is,* $Y \perp B \mid S$ *— then:*

$$I\left(X; Y\right) = I\left(S; Y\right) \tag{4.18}$$

and the information loss due to compression satisfies:

$$\Delta I_q = I\left(S; Y\right) - I\left(\hat{S}; Y\right) \tag{4.19}$$

which depends only on the distortion of the sparse component $\hat{S}$, *not on the total distortion* $D\left(X, \hat{X}\right)$.

The Non-Implication Theorem

Combining Propositions 4.4 and 4.5 yields the central result of this section:

Theorem 4.3 (Non-Implication). *For medical images with sparse diagnostic structure (*$\left|\operatorname{supp}(S)\right| / |\Omega| \ll 1$*):*

$$D(X, \hat{X}) \leq \epsilon \quad \not\Rightarrow \quad D(S, \hat{S}) \leq \epsilon \tag{4.20}$$

and consequently:

$$D(X, \hat{X}) \leq \epsilon \quad \not\Rightarrow \quad \Delta I_q \leq \delta \tag{4.21}$$

for any fixed $\delta > 0$. *That is, bounding global pixel distortion does not bound diagnostic information loss.*

Proof. Consider a compression operator that achieves $D_B = 0$ (perfect background reconstruction) but $D_S = M$ for arbitrarily large M. The global distortion is:

$$D\left(X,\hat{X}\right)=\frac{\left|\text{supp}(S)\right|}{N}\cdot M=0.05\cdot M$$

Setting $\delta=0.05\cdot M$, we have $D(X,\hat{X})=\epsilon$ but $D_S=M=20\epsilon$. The distortion on the diagnostic structure is $20\times$ the global distortion bound. Since ΔI_q depends on D_S, not on $D\left(X,\hat{X}\right)$, the information loss is unbounded by the global distortion.

Empirical Confirmation

The Non-Implication Theorem is directly confirmed by the Part III data. The diagnostic variance ratio $\rho_{\text{diag}}=\left|\text{supp}(S)\right|\cdot\text{Var}(S)/\left(N\cdot\text{Var}(X)\right)\approx 0.04$ (Part II, Section 2.2) quantifies the dominance of background in global statistics.

At Classical SYMPES level L3:

Global MSE is moderate: $\text{PSNR}=27.1$ dB

Lesion MSE is severe: Lesion $\text{PSNR}=22.56$ dB ($\Delta=4.54$ dB below global)

Information loss is catastrophic: $\text{IRR}_{\text{MI}}=0.188$

The gap between global PSNR (27.1 dB) and lesion PSNR (22.56 dB) is exactly what Proposition 4.4 predicts: background distortion dominates global metrics, masking the severe degradation in the diagnostic region. The collapse of IRR_{MI} to 0.188 is exactly what Proposition 4.5 predicts: information loss depends on D_S, not on $D\left(X,\hat{X}\right)$.

Implications for Compression System Design

The Structural Information Model yields a clear design principle: compression systems for medical imaging must allocate bits proportionally not to variance but to diagnostic importance. The classical bit allocation rule — which assigns bits proportionally to σ_i^2 for the i-th component — is optimal for MSE but suboptimal for information preservation.

The optimal allocation under the Structural Information Model would assign bits proportionally to $L_f^{(i)}\cdot\sigma_i^2$, where $L_f^{(i)}$ is the local Lipschitz constant of the clinical feature extractor. This weighted allocation preferentially preserves the sparse diagnostic structure S at the expense of the background B, achieving lower information loss ΔI_q for the same total rate R.

This principle — structure-aware bit allocation — is the operational consequence of the Rate–Structure Paradigm and provides a concrete direction for the design of next-generation medical image compression systems.

Chapter Summary and Transition

Part IV has developed a rigorous mathematical framework for understanding why classical compression theory is structurally insufficient for medical imaging. The principal theoretical contributions are as follows.

First, the task-conditioned rate–distortion formulation (Section 4.1) extends Shannon's framework by replacing pixel distortion with clinical task loss, demonstrating that the resulting optimization landscape loses the convexity properties upon which classical methods depend.

Second, the Fidelity–Topology Decoupling Theorem (Section 4.3, Theorem 4.1) proves constructively — using real data from the TCGA-LUAD cohort — that pixel fidelity and topological fidelity are not monotonically related. Higher PSNR does not guarantee lower topological distortion or higher diagnostic information retention.

Third, the Rate–Topology Bound (Section 4.4, Theorem 4.2) establishes that the topological deviation function $T(R)$ is non-convex, implying that phase transitions in structural quality are an inherent mathematical property of lossy compression, not an algorithmic deficiency.

Fourth, the multi-criteria Clinical Safety condition (Section 4.6) formalizes the conjunctive nature of clinical data integrity, requiring simultaneous satisfaction of energy fidelity, manifold convergence, and topological invariance.

Fifth, the cross-paradigm validation (Section 4.7) demonstrates that the Fidelity–Topology Decoupling holds across two fundamentally different compression architectures (dictionary-based SYMPES and DCT-based TC-SYMPES), confirming that it is a property of the compression problem rather than of any specific algorithm.

Sixth, the Structural Information Model (Section 4.9) provides the definitive mathematical explanation for the Variance–Information Fallacy. The sparse-structure decomposition $X = B + S$ with $|\mathrm{supp}(S)|/|\Omega| \ll 1$ proves that global pixel distortion is dominated by the background component (Proposition 4.4), while clinical information resides entirely in the sparse diagnostic structure (Proposition 4.5). The Non-Implication Theorem (Theorem 4.3) establishes that $D(X, \hat{X}) \le \epsilon \not\Rightarrow \Delta I_q \le \delta$ bounding global distortion does not bound diagnostic information loss. The Information–Distortion Inequality (Proposition 4.3) further shows that the amplification factor $L_f^{\mathrm{ROI}} \gg 1$ explains why moderate pixel distortion produces catastrophic information collapse in diagnostically critical regions.

Part V therefore addresses the implications of these findings for the emerging field of clinical artificial intelligence, where compressed images serve as inputs to deep learning systems. If compression degrades the structural and topological properties upon which AI systems implicitly rely, the consequences for diagnostic safety may be severe.

These findings fundamentally redefine the notion of "acceptable compression" in medical imaging. However, modern clinical decision-making is increasingly mediated by artificial intelligence systems operating in high-dimensional feature spaces. This raises a critical question:

How does compression-induced structural perturbation propagate through learned feature representations and affect AI-driven clinical inference?

The next chapter addresses this question by modelling compression as a perturbation operator acting on feature manifolds, thereby bridging the gap between signal processing theory and clinical AI systems.

CHAPTER 5

Compression in the Age of Clinical AI

Chapter Thesis

Compression acts as a perturbation operator within clinical AI pipelines.

5.0 From Structural Fidelity to Algorithmic Reliability

Part IV demonstrated that pixel fidelity, structural integrity, and diagnostic information preservation are fundamentally distinct quantities. The Fidelity–Topology Decoupling Theorem established that compression may preserve pixel-level distortion bounds while simultaneously altering the structural topology of anatomical representations.

In classical radiology workflows this discrepancy may remain partially concealed. Radiologists interpret images through contextual reasoning, anatomical priors, and clinical experience. Small structural perturbations introduced by compression may therefore be implicitly corrected during visual interpretation.

The emergence of clinical artificial intelligence fundamentally alters this situation. Recent evidence demonstrates that deep learning models can exhibit instabilities in image reconstruction tasks, where small input perturbations produce disproportionate output changes (Antun et al., 2020).

Machine learning systems operate not on raw visual perception but on high-dimensional feature representations extracted from images. These representations encode subtle statistical patterns in intensity distributions, texture structures, and spatial relationships. When compression artifacts modify these patterns, the perturbation propagates directly into the feature space used for algorithmic decision making.

Compression therefore becomes part of the diagnostic pipeline itself. The problem of medical image compression must consequently be reinterpreted: not as a purely signal-processing task but as a transformation applied within a computational diagnostic system.

5.1 Medical Imaging Pipelines in the Era of Clinical AI

Over the past decade, medical imaging has undergone a profound transformation. Traditionally, radiological interpretation relied primarily on visual assessment by trained clinicians. Contemporary medical imaging pipelines, however, increasingly integrate computational analysis and machine-learning–based decision support systems (Hosny et al., 2018).

In modern clinical environments, medical images rarely remain confined to human interpretation alone. Instead, they pass through a complex sequence of computational stages before contributing to diagnostic or prognostic decisions. A simplified representation of a typical imaging pipeline can be expressed as:

Acquisition → Reconstruction → Compression → Feature Extraction → AI Model → Clinical Decision.

Within this pipeline, compression is often treated as a technical necessity rather than a representational transformation. Large-scale imaging repositories, hospital archiving systems, and telemedicine infrastructures routinely rely on compression to reduce storage requirements and facilitate efficient transmission of imaging data (Clark et al., 2013).

Under this operational perspective, compression is assumed to be largely benign as long as global image quality remains visually acceptable and conventional fidelity metrics—such as PSNR or SSIM—remain within acceptable limits.

However, the theoretical and empirical analyzes presented in the preceding chapters suggest that this assumption may be incomplete.

Chapter 2 demonstrated that classical compression frameworks allocate representational resources primarily according to variance distribution. Chapter 3 provided empirical evidence that such variance-preserving compression may nevertheless degrade structural descriptors and topological organization within diagnostically relevant regions (Gökbay, 2013). Chapter 4 then introduced the concept of structural information and argued that the informational content of medical images resides primarily in the stability of anatomical structures rather than in the variance of pixel intensities.

These insights have important implications for modern medical imaging pipelines. If compression alters the structural representation of an image—even subtly—then the features extracted from that image may also change. Because many clinical AI systems rely on feature-based representations derived from imaging data, such structural perturbations may propagate downstream into machine-learning models.

In other words, compression does not merely reduce file size. It may also influence the feature space on which clinical AI systems operate.

This observation reframes the role of compression in medical imaging. Instead of functioning solely as a storage optimization technique, compression becomes part of the representational transformation that shapes the input to computational diagnostic systems.

The implications are particularly important in the context of radiomics and deep-learning–based imaging biomarkers. Radiomic pipelines extract quantitative descriptors—such as texture statistics, shape features, and intensity distributions—from medical images (Lambin et al., 2012; Gillies et al., 2016). These descriptors are then used to train predictive models for tasks such as tumor classification, treatment response prediction, and survival estimation (Aerts et al., 2014).

If compression modifies the structural organization of the image, even subtly, the extracted radiomic features may change accordingly. Such changes may shift the position of a sample within feature space, potentially altering the predictions generated by machine-learning models.

The relationship between compression and AI performance therefore raises a critical question: to what extent do compression-induced structural perturbations influence the reliability of clinical AI systems?

Addressing this question requires examining how compression interacts with the feature extraction processes that underpin modern medical imaging analytics. The following sections analyze this interaction by examining representation bias, feature instability, and the propagation of structural perturbations through AI pipelines.

5.2 Compression as a Perturbation of Feature Manifolds

The impact of image compression on medical artificial intelligence cannot be adequately understood at the level of pixel distortion alone. Classical fidelity measures such as mean squared error (MSE), peak signal-to-noise ratio (PSNR), and structural similarity index (SSIM) evaluate the deviation between the original image x and its compressed version $\tilde{x}$ in the observation space.

However, contemporary clinical AI systems do not operate directly on raw pixel intensities in a diagnostically meaningful sense. Rather, they transform images into hierarchical internal representations, or feature manifolds, on which downstream classification, segmentation, detection, survival prediction, or treatment response modeling is performed. From this perspective, compression must be treated not merely as a signal-processing operation but as a geometric perturbation of the learned representation space.

Let $f_\theta : X \to Z$ denote a feature extraction map parameterized by a model θ, where X is the image space and $Z \subseteq \mathbb{R}^d$ is the latent feature space. For an original medical image $x \in X$, the corresponding latent representation is $z = f_\theta(x)$, while for its compressed counterpart $x^{\sim} = C_\lambda(x)$, obtained under compression operator Cλ with compression level λ, the latent code becomes $\tilde{z} = f_\theta(\tilde{x}) = f_\theta(C_\lambda(x))$.

The central question is therefore not simply whether x and $\tilde{x}$ are visually similar, but whether z and $\tilde{z}$ remain sufficiently close in the representation space such that the clinical inference function $g : Z \to Y$ preserves the diagnostic decision: $g(z) \approx g(\tilde{z})$.

This reframing is crucial. In clinical AI, the diagnostically relevant information is not uniformly distributed across the image domain. It is concentrated in sparse anatomical boundaries, subtle texture gradients, topological relations, contrast wash-in and wash-out dynamics, lesion margins, or regional intensity patterns whose statistical footprint may be small in the pixel domain but disproportionately large in the latent decision geometry. A compression scheme that preserves global pixel fidelity while perturbing these task-relevant structures may leave PSNR nearly intact yet induce a substantial displacement in the learned feature manifold. In other words, small distortion in image space does not imply small distortion in feature space.

5.2.1 Latent Representation Geometry Under Compression

Deep neural networks trained on medical images implicitly construct structured manifolds in latent space. Samples belonging to similar anatomical or pathological classes tend to cluster, while decision boundaries separate clinically distinct groups. Suppose that the model's diagnostic decision is based on a latent margin $m(x) = h(f_\theta(x))$, where h is a classifier head or task-specific predictor. Compression alters this margin through a perturbation term $\Delta z = \tilde{z} - z = f_\theta(C_\lambda(x)) - f_\theta(x)$.

If fθ is locally Lipschitz-continuous around x, then $\Delta z \leq L\, C_\lambda(x) - x$, for some local Lipschitz constant $L > 0$. At first glance, this may seem reassuring: small image perturbations should produce bounded feature perturbations. But

the devil, that elegant little saboteur, lives in the geometry. In high-dimensional nonlinear networks, local sensitivity is highly anisotropic. Compression artifacts are not random isotropic noise; they are structured perturbations aligned with transform bases, quantization bins, interpolation kernels, denoising priors, or codec-specific regularization mechanisms. As a result, the mapping from pixel disturbance to latent displacement is not governed merely by magnitude, but by direction relative to the manifold curvature and decision boundary orientation.

Thus, even when $\|C\lambda(x) - x\|$ is small, the induced $\|\Delta z\|$ may be disproportionately large if the perturbation projects onto diagnostically sensitive feature directions. Formally, for local linearization, $\Delta z \approx J_f(x)\Delta x$, where $J_f(x)$ is the Jacobian of f_θ at x, and $\Delta x = C_\lambda(x) - x$. The effect of compression therefore depends on the interaction between the artifact pattern and the Jacobian spectrum of the feature extractor. If Δx aligns with high-gain singular directions of $Jf(x)$, then clinically relevant latent variables may shift significantly despite modest pixel error. This is precisely why compression can be benign for visual interpretation yet harmful for machine inference.

5.2.2 Manifold Drift Versus Pixel Fidelity

The distinction between pixel fidelity and manifold fidelity is especially important in medical imaging because many diagnostically critical cues are weak, localized, and structurally constrained. Consider a lesion classification task in thoracic CT, breast DCE-MRI, or liver lesion characterization. The relevant signal may reside in subtle spiculations, edge irregularity, heterogeneous enhancement, or fine texture distributions. These features contribute little to global MSE, yet they may anchor the sample's position within a class-discriminative manifold. Compression that smooths, attenuates, or discretizes these subtle structures can cause manifold drift, meaning that the compressed sample moves away from its native class neighborhood and toward another region of latent space.

Let $M_k \subset Z$ denote the latent manifold associated with class k. For a sample x with true label $y = k$ reliable inference requires that $z = f_\theta(x)$ remain close to Mk and sufficiently distant from competing manifolds Mj, $j \neq k$. Compression introduces risk whenever $d(\tilde{z}, M_k) > d(z, M_k)$ or more severely when $d(\tilde{z}, M_j) < d(\tilde{z}, M_k)$ for some $j \neq k$.

In such cases, the compressed representation effectively crosses a diagnostic margin. Importantly, this failure can occur without conspicuous visual degradation. The radiologist may still judge the image acceptable, while the AI model silently experiences a representation-level shift that degrades

prediction reliability. This asymmetry between human-perceived quality and machine-relevant stability is a central thesis of this monograph.

5.2.3 Compression as Structured, Not Random, Perturbation

A further conceptual error in much of the literature is the implicit treatment of compression artifacts as if they were generic noise. They are not. Compression acts through structured operators—transform truncation, coefficient quantization, entropy coding, motion compensation, prediction residual modeling, learned bottleneck regularization, or patch-token discretization in foundation-model preprocessing pipelines. Consequently, the perturbation Δx is highly non-random and often correlated with the very structures the AI system relies upon.

For example, block-based codecs may introduce local discontinuities at patch boundaries; transform-based quantization may suppress high-frequency textural detail; denoising-oriented learned compression may remove low-amplitude heterogeneity mistaken for nuisance variation; temporal or multi-phase compression may alter enhancement kinetics in dynamic imaging. Such perturbations are semantically entangled with clinically meaningful content. Therefore, compression should be modeled as a biased transformation of the data-generating process, not as an innocuous storage convenience.

This distinction matters because modern learning theory shows that model robustness depends not only on noise magnitude but also on whether perturbations are on-manifold, near-manifold, or adversarially off-manifold. Compression artifacts may shift data into regions of feature space that were sparsely represented during training, effectively behaving as a domain perturbation. In this sense, compression occupies an intermediate territory between classical corruption and distribution shift: it begins as a deterministic signal transform but propagates downstream as a representation-level covariate deformation.

5.2.4 Clinical AI as a Composition of Operators

A useful formalization is to treat the end-to-end medical AI pipeline as a composition of operators: $x \to C_\lambda \to \tilde{x} \to P \to \hat{x} \to f_\theta \to z \to g \to \hat{y}$, where $C\lambda$ is compression, P may include decompression, resampling, normalization, windowing, denoising, registration, or segmentation pre-processing, $f\theta$ is the learned representation map, and g is the downstream clinical predictor. Under this view, compression is not external to inference; it is part of the effective operator chain that determines the clinical output. The final decision is therefore governed by $\hat{y} = g\left(f_\theta\left(P\left(C_\lambda\left(x\right)\right)\right)\right)$.

The implication is profound: if compression is present during deployment but absent, mismatched, or insufficiently represented during model development, then the learned decision function is evaluated on an altered representation manifold. This mismatch creates a hidden source of error that cannot be diagnosed through standard image-quality metrics alone. In other words, compression becomes a latent confounder in the AI pipeline.

5.2.5 Representation Sensitivity and Task Dependence

Not all clinical AI tasks are equally vulnerable to compression-induced manifold perturbation. Sensitivity depends on at least four interacting factors.

First, it depends on the task granularity. Coarse tasks such as gross organ localization may tolerate moderate manifold drift, whereas fine-grained tasks such as micro-lesion detection, margin characterization, radiomics-based risk prediction, or longitudinal response tracking may be acutely sensitive.

Second, it depends on the feature regime. Models relying on deep texture statistics, radiomic heterogeneity, or topological descriptors are more likely to suffer when compression suppresses high-frequency or structural information. Conversely, tasks driven by large-scale shape cues may be relatively more robust.

Third, it depends on the distance to decision boundaries. Samples near classification margins are inherently fragile; even small latent displacement can flip the prediction. A compression level that is harmless for confidently classified cases may therefore be dangerous for borderline or rare phenotypes.

Fourth, it depends on the training distribution. If the model has encountered compression-consistent examples during training, some invariance may emerge. If not, the same perturbation may produce severe out-of-distribution behavior. Thus, representation stability under compression is not an intrinsic property of the image alone, but a joint property of the codec, the task, the model, and the data distribution.

5.2.6 Toward Quantifying Manifold Perturbation

To rigorously evaluate compression in the era of clinical AI, one must move beyond PSNR/SSIM and quantify representation-level deformation directly. Let a batch of original images $\{x_i\}_{i=1}^{N}$ and their compressed versions $\{\tilde{x}_i\}_{i=1}^{N}$ be given. Several quantities become relevant:

The average latent displacement: $D_{\text{latent}} = \frac{1}{N}\sum_i f_\theta(x_i) - f_\theta(\tilde{x}_i)$.

The class-center deviation: $D_{\text{class}} = \frac{1}{N}\sum_i \left(f_\theta(\tilde{x}_i) - \mu_{y_i}\right) - \left(f_\theta(x_i) - \mu_{y_i}\right)$, where μ_{y_i} of class y_i in feature space.

The local neighborhood preservation rate, measuring whether compressed samples retain their nearest neighbors in latent space.

The topological persistence of the feature cloud, evaluated through graph connectivity, Betti numbers, or persistent homology summaries before and after compression.

The predictive consistency: $R_{\text{pred}} = \frac{1}{N}\sum_i 1\left[g\left(f_\theta\left(x_i\right)\right) = g\left(f_\theta\left(\tilde{x}_i\right)\right)\right]$.

These measures are not merely technical embellishments; they reflect a deeper epistemic shift. The question is no longer "How much does compression distort the image?" but "How much does compression distort the inferential geometry on which clinical AI depends?"

5.2.7 Conceptual Consequence

The conceptual consequence of this section is straightforward but far-reaching. Compression in medical imaging should not be regarded as a neutral storage layer preceding analysis. Once AI enters the pipeline, compression becomes a representation-shaping operator. It perturbs the geometry of the feature manifold, alters local class neighborhoods, changes the margin structure of decision functions, and may silently compromise downstream reliability even when conventional image-quality measures remain favorable.

Accordingly, the assessment of medical image compression must be elevated from visual fidelity analysis to algorithmic reliability analysis. This transition marks a decisive break from the classical variance-preservation doctrine developed in earlier chapters. Pixel-level similarity may preserve appearance, but it does not guarantee preservation of the latent structures through which clinical AI reasons. In the age of clinical machine learning, the relevant unit of fidelity is no longer the pixel alone; it is the stability of diagnostically meaningful representation.

If compression perturbs the geometry of feature manifolds, then the next question is no longer whether representations move, but whether they remain stable enough to preserve clinically meaningful information across tasks, models, and operating conditions. This motivates the analysis of feature instability and information retention in the following section.

5.3 Feature Instability and Information Retention

The previous section established that image compression can perturb the geometry of learned feature manifolds in clinical artificial intelligence systems. However, manifold perturbation alone does not necessarily imply a loss of clinically relevant information. The critical question is whether the

compressed representation preserves the diagnostic signal embedded in the original image. In other words, the fundamental issue is not merely feature displacement, but feature instability with respect to the information required for reliable inference.

In contemporary medical AI systems, images are mapped into latent feature vectors that serve as sufficient statistics for downstream tasks. Let the feature extraction function be defined as $z = f_\theta(x)$, where x denotes the original medical image and $z \in \mathbb{R}^d$ represents the learned latent representation. Under compression operator Cλ, the observed image becomes $\tilde{x} = C_\lambda(x)$, leading to a perturbed feature representation $\tilde{z} = f_\theta(\tilde{x})$.

The central problem of this section can therefore be formulated as the information preservation condition $I(Y;z) \approx I(Y;\tilde{z})$, where Y denotes the clinically relevant variable (diagnosis, segmentation mask, survival outcome, or treatment response). If this equality fails, compression has effectively removed or distorted information necessary for clinical inference.

Importantly, classical compression metrics do not evaluate this condition. Measures such as PSNR or SSIM quantify similarity between x and $\tilde{x}$, but they do not measure the mutual information between the latent representation and the diagnostic variable. As a result, compression may appear acceptable from an image-quality perspective while simultaneously degrading the predictive content of the representation.

5.3.1 Feature Stability as a Diagnostic Requirement

For clinical AI systems to remain reliable under compression, the representation mapping must exhibit feature stability, meaning that diagnostically relevant coordinates in feature space remain invariant or nearly invariant under the compression transformation.

Formally, consider a feature vector z consisting of components $z = (z_1, z_2, \ldots, zd)$. Some components correspond to nuisance variation (scanner noise, acquisition variability, irrelevant background patterns), while others encode diagnostically meaningful information. Let $S \subset \{1,\ldots,d\}$ denote the set of diagnostically relevant feature indices.

Feature stability requires that for all $k \in S$ $|z_k - \tilde{z}_k| \leq \varepsilon$, for some small tolerance parameter ε. Violation of this condition implies that compression has altered the representation of clinically meaningful structures.

The difficulty arises because modern deep learning models do not explicitly disentangle these feature roles. Instead, representations emerge implicitly through training, meaning that compression may affect latent variables in ways

that are difficult to interpret. A perturbation that appears minor in magnitude may nevertheless disrupt a feature that strongly influences the final prediction.

5.3.2 Information Retention Under Representation Perturbation

A principled way to analyze this phenomenon is through the lens of information retention. Let the original representation contain diagnostic information quantified by $I(Y;z)$. After compression, the available information becomes $I(Y;\tilde{z})$. The information retention ratio can therefore be defined as $R_{info}=\frac{I(Y;\tilde{z})}{I(Y;z)}$.

A value close to one indicates that compression preserves most diagnostically relevant information, whereas a significant decrease suggests that the representation has lost predictive content.

This concept highlights a fundamental tension between rate–distortion optimization and diagnostic information preservation. Classical compression theory minimizes expected distortion $D=E\left[x-\tilde{x}^2\right]$, subject to a bitrate constraint. However, diagnostic tasks are not sensitive to global pixel distortion per se. Instead, they depend on the preservation of task-relevant structural information. Consequently, minimizing distortion does not guarantee maximal $I(Y;\tilde{z})$. In fact, compression strategies that aggressively remove high-frequency components or smooth local heterogeneity may inadvertently discard features that carry predictive value.

5.3.3 Sensitivity of Clinical AI Models to Representation Perturbation

The degree to which compression-induced feature instability affects prediction depends strongly on the structure of the downstream model. Consider a classifier $\hat{\mathbf{y}}=g(z)$, where g is typically implemented as a fully connected network, transformer head, or logistic regression layer.

Under representation perturbation $z\to\tilde{z}=z+\Delta z$, the prediction becomes $g(\tilde{z})=g(z+\Delta z)$. Using first-order approximation, $g(z+\Delta z)\approx g(z)+\nabla g(z)\cdot\Delta z\cdot$

Thus, prediction stability depends on both the magnitude of the feature perturbation and the local gradient of the classifier. If the gradient is large along the direction of Δz, even a small compression-induced displacement may cause a significant change in prediction probability.

This observation explains why compression effects are often case dependent. Samples located near decision boundaries are particularly vulnerable because their prediction margins are small. For such cases, minimal feature instability

can result in label inversion or large probability shifts. Conversely, samples deep within class clusters may remain robust even under moderate compression.

5.3.4 Instability in Radiomics and Texture-Based Features

Feature instability is especially critical for radiomics-based systems, where quantitative descriptors capture subtle statistical patterns within lesions or tissue regions (Lambin et al., 2012; Gillies et al., 2016). Radiomic features often include gray-level co-occurrence statistics, texture entropy, local variance patterns, fractal or morphological descriptors, and wavelet-domain features.

These features are inherently sensitive to spatial discretization, quantization, and smoothing operations. Compression artifacts that modify pixel correlations or attenuate high-frequency components can therefore cause significant variation in radiomic descriptors.

Empirical studies in lung CT, breast MRI, and liver imaging have demonstrated that even moderate compression ratios can alter radiomic feature values beyond acceptable reproducibility thresholds (Aerts et al., 2014). Such instability propagates through machine learning models trained on these descriptors, potentially altering risk stratification or treatment response predictions.

This vulnerability illustrates an important principle: diagnostic features often occupy precisely the spectral bands that compression algorithms attempt to suppress.

5.3.5 Representation Drift and Predictive Consistency

An operational way to evaluate feature stability is through predictive consistency, defined as the probability that compressed and uncompressed images yield identical predictions: $R_{pred} = P\left(g\left(f\theta(x)\right) = g\left(f\theta(\tilde{x})\right)\right)$.

High predictive consistency indicates that compression does not materially affect the AI system's decision process. However, consistency alone is not sufficient; a model may remain consistently wrong under compression. Therefore predictive consistency must be interpreted jointly with diagnostic accuracy and calibration.

Another useful concept is representation drift, defined as the average feature displacement $D_{feat} = E\left[\| f\theta(x) - f\theta(\tilde{x}) \|\right]$. Empirical analysis often reveals a nonlinear relationship between pixel distortion and representation drift. Small increases in pixel error may correspond to disproportionately large shifts in feature space, particularly when compression artifacts interact with sensitive representation directions learned by the model.

5.3.6 Compression as Selective Information Filtering

Compression can therefore be interpreted as a selective information filter applied prior to AI inference. Transform coding, coefficient quantization, and learned bottleneck compression mechanisms effectively decide which components of the signal are preserved and which are discarded.

From a signal-processing perspective this filtering is designed to remove perceptually redundant information. From a clinical AI perspective, however, perceptual redundancy does not necessarily correspond to diagnostic irrelevance. Features invisible or negligible to the human eye may nonetheless carry statistical patterns that a machine learning model exploits.

This mismatch between human perceptual importance and machine-learned predictive importance lies at the heart of compression-induced information loss in AI pipelines.

5.3.7 Conceptual Implication

The analysis in this section reveals a critical conceptual shift. In classical imaging workflows, compression was evaluated primarily with respect to visual interpretability by human observers. In modern AI-driven clinical systems, the relevant criterion becomes representation stability with respect to the diagnostic task.

Compression must therefore be evaluated not only in the pixel domain but also in the information domain of the learned representation. Failure to do so risks silently degrading the predictive content of medical images while leaving conventional quality metrics largely unaffected.

In this sense, compression does not merely distort the signal; it selectively reshapes the information available to the clinical AI model. Understanding and quantifying this transformation is essential for ensuring reliable deployment of machine learning systems in medical imaging.

5.4 Compression-Induced Distribution Shift

The preceding sections established that compression can perturb the geometry of feature manifolds and destabilize diagnostically relevant representations. While these effects may initially appear as instance-level perturbations, their consequences extend beyond individual samples. When compression systematically alters the statistical structure of input images, it induces a distribution shift between the data used during model development and the data encountered during clinical deployment.

This phenomenon is particularly important in medical artificial intelligence because machine learning models implicitly assume that training and deployment data are drawn from the same underlying distribution. When this assumption is violated, predictive performance may degrade even if the model architecture and training procedure remain unchanged.

Let X denote the random variable representing medical images and Y the corresponding clinical label. A learning algorithm estimates a decision function $\hat{y} = g\left(f\theta\left(x\right)\right)$ based on samples drawn from a training distribution $\left(x,y\right) \sim P_{train}\left(X,Y\right)$. During deployment, however, images may undergo compression before entering the inference pipeline: $\tilde{x} = C\lambda\left(x\right)$. The deployed model therefore operates on samples drawn from a different distribution $\left(\tilde{x},y\right) \sim P_{deploy}\left(X,Y\right)$. If compression modifies the statistical structure of images, then $P_{deploy}\left(X,Y\right) \neq P_{train}\left(X,Y\right)$. This discrepancy constitutes a covariate shift, in which the marginal distribution of the input changes while the underlying label semantics remain constant.

5.4.1 Compression as a Deterministic Distribution Transformer

Unlike many sources of dataset shift, compression is not random. It is a deterministic operator applied systematically across the dataset. This property has two important consequences.

First, the induced distribution shift is structured rather than stochastic. Transform-based compression suppresses high-frequency components, introduces quantization artifacts, and modifies local intensity correlations. These transformations reshape the statistical properties of the dataset in a predictable but nontrivial way.

Second, because compression is applied uniformly to all images, the resulting dataset may exhibit consistent but altered feature distributions. For example, histogram statistics, texture descriptors, and high-order radiomic features may systematically shift after compression. Even if these changes are subtle at the pixel level, they can accumulate across features and lead to measurable displacement in the learned representation space.

Consequently, the dataset used during inference may lie on a different manifold than the dataset used during training.

5.4.2 Representation-Level Distribution Shift

Distribution shift becomes more evident when viewed in the latent representation space. Let $z = f\theta\left(x\right)$ denote the feature representation of

the original image and $\tilde{z} = f\theta\left(C\lambda\left(x\right)\right)$ the representation of the compressed image.

If compression consistently perturbs representations, the empirical feature distribution changes from $P_{train}\left(Z\right)$ to $P_{deploy}\left(Z\right)$. Even if this shift is modest in magnitude, it may alter the geometry of class clusters, distort decision boundaries, and affect calibration of predictive probabilities.

Such effects are well known in machine learning theory: classifiers trained under one distribution may perform unpredictably when evaluated under another (Zech et al., 2018). In medical AI, where decision boundaries often depend on subtle imaging patterns, even small distribution shifts can produce clinically meaningful errors.

5.4.3 Relationship to Domain Shift and Dataset Bias

Compression-induced distribution shift is conceptually related to other forms of domain shift observed in medical imaging, including scanner variability, reconstruction kernel differences, acquisition protocol changes, and demographic dataset imbalance.

However, compression differs from these factors in an important way. While scanner differences arise from hardware or acquisition protocols, compression acts as a post-acquisition transformation that modifies the signal before analysis. As a result, it may interact with downstream processing steps such as normalization, resampling, or feature extraction in complex ways.

In effect, compression can amplify existing dataset biases by introducing additional transformation layers between acquisition and analysis.

5.4.4 Implications for Model Robustness

Distribution shift induced by compression presents a challenge for model robustness. Standard evaluation procedures often assume that training, validation, and test datasets share similar preprocessing pipelines. If compression is applied inconsistently across these datasets, evaluation metrics may provide an overly optimistic estimate of real-world performance.

For example, a model trained and validated on uncompressed images may demonstrate excellent accuracy during development. Yet when deployed in clinical environments where images are stored in compressed formats, the same model may encounter a shifted input distribution and experience degraded performance.

This problem is rarely detected during model development because compression is often treated as a storage issue rather than an analytical variable.

5.4.5 Compression and Calibration Drift

Distribution shift does not only affect classification accuracy. It may also alter the calibration of predictive probabilities. Calibration refers to the alignment between predicted probabilities and observed outcomes. When the input distribution changes, the relationship between feature representations and labels may shift, leading to systematic overconfidence or underconfidence in model predictions.

In clinical contexts, calibration errors can be as dangerous as misclassifications. Overconfident incorrect predictions may mislead clinicians, while underconfident predictions may reduce trust in the AI system.

Therefore, evaluating compression effects requires examining both accuracy and calibration stability.

5.5 Compression as a Hidden Component of Clinical AI

The preceding sections demonstrate that compression can perturb feature manifolds, destabilize diagnostically relevant representations, and induce dataset-level distribution shifts. These observations lead to a crucial conceptual conclusion: compression is not merely a storage mechanism but an operational component of the clinical AI pipeline.

Traditional imaging workflows treat compression as an external preprocessing step whose purpose is to reduce storage and transmission costs. In contrast, AI-based diagnostic systems implicitly incorporate compression into the effective transformation applied to the input signal.

Consider the end-to-end inference pipeline: $x \to C_\lambda(x) \to P\left(C_\lambda(x)\right) \to f_\theta\left(P\left(C_\lambda(x)\right)\right) \to g\left(f_\theta(\cdot)\right)$. Here compression $C\lambda$ appears as the first operator in the processing chain. Because downstream components operate on the compressed signal rather than the original image, compression directly influences the information available to the model.

5.5.1 The Hidden Operator Problem

Despite its influence, compression is rarely modeled explicitly in machine learning pipelines. Instead, it is often treated as an implementation detail within hospital imaging infrastructure. This oversight creates what may be called the hidden operator problem.

A hidden operator is a transformation that affects the data distribution but is not explicitly accounted for during model development or evaluation. Compression fits this definition because it modifies the input signal while remaining invisible in the conceptual design of the AI system.

As a result, the learned model implicitly depends on the properties of the compression operator, even if the training procedure never considered it explicitly.

5.5.2 Training-Deployment Mismatch

The hidden operator problem becomes particularly severe when compression conditions differ between training and deployment environments. Training pipeline: $x \rightarrow f_\theta(x)$. Deployment pipeline: $x \rightarrow C_\lambda(x) \rightarrow f_\theta(C_\lambda(x))$

The learned representation is therefore optimized for one input distribution but evaluated on another. In extreme cases, this mismatch may produce systematic performance degradation, especially for tasks relying on subtle structural or texture features.

From a systems perspective, this mismatch represents an uncontrolled variable within the AI pipeline.

5.5.3 Compression in the Context of Regulatory AI

Clinical AI systems increasingly operate within regulatory frameworks that require reproducibility, traceability, and validation. If compression affects diagnostic predictions, it must be considered part of the system under validation.

Ignoring compression during validation risks underestimating the variability of model outputs in real-world deployment environments. Consequently, compression should be documented and controlled as part of the technical specification of clinical AI systems.

5.6 Implications for Clinical Deployment

The presence of compression within clinical AI pipelines has practical implications for the safe deployment of machine learning systems in medical imaging environments.

First, AI models must be evaluated under realistic data conditions that reflect the compression formats used in clinical storage and transmission systems. Testing models solely on uncompressed datasets may fail to capture performance degradation occurring in routine clinical workflows.

Second, compression parameters should be treated as part of the data preprocessing specification of the AI system. Just as acquisition protocols and reconstruction kernels are documented in radiological practice, compression settings may need to be standardized to ensure consistent model behavior.

Third, developers should assess model robustness to compression during training. This may involve including compressed images in the training dataset or designing models that remain stable across a range of compression levels.

Finally, clinicians and system integrators must recognize that storage decisions can influence AI performance. Choices made for cost efficiency or bandwidth optimization may inadvertently affect diagnostic reliability.

5.6.1 Toward Compression-Aware AI Evaluation

To mitigate these risks, evaluation frameworks should incorporate compression as a controlled experimental factor. Benchmark datasets can be generated at multiple compression levels, allowing systematic assessment of model performance under varying storage conditions.

Key evaluation metrics may include predictive accuracy across compression ratios, representation stability, predictive consistency, and calibration robustness. Such analyzes would provide a more realistic estimate of how AI systems behave in operational environments.

5.6.2 Bridging the Gap Between Imaging Infrastructure and AI

Ultimately, compression highlights a broader issue in medical AI: the disconnect between imaging infrastructure and algorithm development. Radiology workflows involve multiple transformation stages—reconstruction, normalization, compression, archiving, and transmission—yet many AI studies treat images as static inputs detached from this pipeline.

A comprehensive understanding of clinical AI requires integrating these stages into a unified analytical framework. Compression, as a ubiquitous yet often overlooked transformation, represents a critical step toward this integration.

5.7 Toward Task-Aware Medical Image Compression

The analysis developed throughout this chapter reveals a fundamental limitation of classical image compression in the context of clinical artificial intelligence. Traditional compression algorithms are designed to minimize perceptual or pixel-level distortion under bitrate constraints. However, clinical

AI systems do not operate on pixels alone; they rely on learned representations that encode diagnostically relevant structures within complex feature manifolds.

Consequently, compression optimized for visual fidelity does not necessarily preserve the information required for reliable machine inference. This mismatch motivates the development of task-aware medical image compression, in which compression strategies are explicitly designed to preserve information relevant to downstream analytical tasks.

Rather than treating compression as an independent storage layer, task-aware compression integrates the objectives of image coding and machine inference into a unified framework.

5.7.1 From Perceptual Fidelity to Task Fidelity

Classical compression theory optimizes a rate–distortion objective of the form $\min_C D(x,\tilde{x})+\lambda R$, where $D(x,\tilde{x})$ measures pixel distortion and R denotes bitrate.

In medical AI applications, however, the relevant objective should consider the effect of compression on the downstream diagnostic task. A more appropriate formulation therefore includes a task-dependent loss: $\min_C D(x,\tilde{x})+\lambda R+\alpha \mathrm{L}_{\text{task}}$, where L_{task} quantifies the degradation of predictive performance caused by compression.

This formulation shifts the optimization goal from preserving visual similarity to preserving task-relevant information. In practical terms, the objective becomes not merely to reconstruct an image that looks similar to the original, but to ensure that the compressed representation leads to the same clinical decision.

5.7.2 Feature-Aware Compression

One approach to task-aware compression is feature-aware encoding, in which compression algorithms prioritize regions or frequency components that contribute most strongly to diagnostic features.

In deep learning pipelines, these regions can be estimated using techniques such as saliency maps, gradient-based attribution, feature activation maps, and attention mechanisms.

By identifying structures that influence the model's prediction, compression algorithms can allocate bitrate preferentially to diagnostically relevant regions while applying stronger compression to less informative areas. This strategy aligns compression decisions with the information structure of the predictive model rather than with global image statistics.

5.7.3 Representation-Preserving Compression

A second approach focuses directly on preserving the stability of learned feature representations. Instead of minimizing pixel distortion, compression can be optimized to minimize representation drift: $\min_C f_\theta(x) - f_\theta(\tilde{x}) f_\theta$ the feature extraction function of the clinical AI model.

This objective ensures that compression does not alter the latent representation on which the diagnostic decision depends. In effect, the compression algorithm becomes aware of the internal geometry of the machine learning model. Such approaches can be implemented using differentiable compression models or neural codecs that allow joint optimization of the compression pipeline and the inference model.

5.7.4 Joint Optimization of Compression and Inference

The most integrated strategy involves joint training of compression and inference systems. In this framework, the compression model and the diagnostic model are trained together to optimize both storage efficiency and predictive performance.

Let E_φ denote an encoder producing a compressed representation y, and D_φ the corresponding decoder. The inference model operates on the reconstructed image: $\hat{y} = g_\theta\left(D_\varphi\left(E_\varphi(x)\right)\right)$. Joint optimization seeks parameters Φ and θ that minimize $L = L_{task}\left(g_\theta\left(D_\varphi\left(E_\varphi(x)\right)\right), y\right) + \beta R(y)$.

This formulation treats compression not as a preprocessing step but as a component of the predictive system itself. In this paradigm, the compressed representation becomes a task-optimized latent code that balances storage efficiency with diagnostic reliability.

5.7.5 Practical Constraints in Clinical Environments

Although task-aware compression offers a promising direction, practical implementation within clinical imaging infrastructure must address several constraints.

First, medical imaging systems are tightly integrated with existing standards such as DICOM and PACS. Compression strategies must therefore remain compatible with established storage and transmission frameworks.

Second, clinical AI models often evolve over time as new datasets and training strategies become available. Compression schemes optimized for a particular model may not generalize across different architectures or tasks.

Third, regulatory requirements demand reproducibility and interpretability in clinical decision-support systems. Compression methods that alter the representation in opaque ways may complicate validation and certification processes.

These constraints suggest that task-aware compression should aim not only for optimal performance but also for interoperability and stability within clinical workflows.

5.7.6 Toward a New Paradigm of Medical Image Coding

The emergence of AI-driven diagnostics suggests a broader conceptual shift in medical image compression. Historically, compression algorithms were designed to serve human observers. In the era of clinical AI, images increasingly serve dual audiences: radiologists and machine learning systems.

This dual role challenges the traditional assumption that perceptual similarity is the appropriate objective for image coding. Instead, compression must consider the informational requirements of both human interpretation and automated analysis.

Task-aware compression therefore represents a step toward a new paradigm of medical image coding in which storage efficiency, visual fidelity, and algorithmic reliability are jointly optimized.

The development of task-aware compression closes the conceptual arc of this chapter. Earlier sections demonstrated that classical compression can perturb feature manifolds, destabilize diagnostic representations, and induce distribution shifts that affect clinical AI systems. Recognizing compression as a hidden component of the AI pipeline leads naturally to the need for compression strategies that explicitly account for downstream analytical tasks. The next section summarizes the key conceptual contributions of this chapter and situates them within the broader argument of the monograph.

Chapter Summary and Transition

This chapter examined the role of image compression within contemporary clinical artificial intelligence pipelines and demonstrated that compression cannot be treated merely as a storage or transmission mechanism when medical images serve as inputs to machine learning systems. Instead, compression acts as a transformation that reshapes the information structure available to downstream algorithms.

The analysis began by reframing compression from a signal-processing operation to a geometric perturbation within representation space. Section 5.2

showed that compression can alter the geometry of learned feature manifolds by introducing structured perturbations into the input signal. Although such perturbations may appear small in pixel space, their effect in latent representation space can be substantially larger because diagnostic models rely on complex nonlinear feature mappings.

Section 5.3 extended this observation by examining the consequences of manifold perturbation for feature stability and information retention. Diagnostic models depend on the preservation of latent features that encode clinically relevant structures. Compression that modifies these features can reduce the mutual information between the representation and the clinical label, thereby degrading the predictive content of the image even when visual similarity remains high.

Section 5.4 demonstrated that the effects of compression are not limited to individual images. When applied systematically across datasets, compression modifies the statistical structure of input data and induces a distribution shift between training and deployment environments. Because machine learning models assume that training and inference data follow similar distributions, such shifts may lead to reduced robustness, degraded accuracy, and calibration drift in clinical AI systems.

Section 5.5 addressed the conceptual implications of these findings by identifying compression as a hidden operator within the clinical AI pipeline. Although compression is typically treated as an infrastructure-level process within imaging systems, it directly influences the signal processed by machine learning models. Consequently, compression becomes an implicit component of the decision-making system and must be considered during model design, evaluation, and validation.

Section 5.6 examined the practical consequences of this hidden operator for clinical deployment. AI systems trained on uncompressed images may encounter compressed data in real-world clinical environments, creating a mismatch between development and deployment conditions. This mismatch can compromise diagnostic reliability unless compression is explicitly incorporated into evaluation protocols.

Finally, Section 5.7 proposed a conceptual direction for addressing this challenge through task-aware medical image compression. Rather than optimizing compression solely for pixel fidelity or perceptual similarity, task-aware approaches aim to preserve diagnostically relevant information within the representations used by machine learning systems. Such strategies integrate

compression with downstream analytical tasks, aligning storage efficiency with algorithmic reliability.

Taken together, the results of this chapter establish that compression is not a neutral preprocessing step in the era of clinical artificial intelligence. By perturbing feature manifolds, destabilizing representations, and altering dataset distributions, compression becomes a factor that directly influences the behavior and reliability of diagnostic algorithms.

These observations reinforce the central thesis of this monograph: pixel-level fidelity is not equivalent to diagnostic information preservation. As medical imaging increasingly intersects with machine learning, evaluating compression solely through pixel-based metrics becomes insufficient. Reliable clinical AI systems require compression strategies that account for the informational structure of diagnostic representations.

In the broader argument of this work, the findings of this chapter bridge the gap between the theoretical critique of variance-based compression developed in earlier chapters and the empirical analyzes presented in Part III. Compression, once considered a peripheral technical detail, emerges as a central variable in the design of trustworthy medical AI systems.

References

Aerts, H. J. W. L., Velazquez, E. R., Leijenaar, R. T. H., Parmar, C., Grossmann, P., Carvalho, S., Bussink, J., Monshouwer, R., Haibe-Kains, B., Rietveld, D., Hoebers, F., Rietbergen, M. M., Leemans, C. R., Dekker, A., Quackenbush, J., Gillies, R. J., & Lambin, P. (2014). Decoding tumor phenotype by noninvasive imaging using a quantitative radiomics approach. Nature Communications, 5, 4006. https://doi.org/10.1038/ncomms5006

Antun, V., Renna, F., Poon, C., Adcock, B., & Hansen, A. C. (2020). On instabilities of deep learning in image reconstruction and the potential costs of AI. Proceedings of the National Academy of Sciences, 117(48), 30088–30095. https://doi.org/10.1073/pnas.1907377117

Bankman, I. (Ed.). (2008). Handbook of medical image processing and analysis. Elsevier.

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. Journal of the Royal Statistical Society: Series B (Methodological), 57(1), 289–300. https://doi.org/10.1111/j.2517-6161.1995.tb02031.x

Berenguer, R., Pastor-Juan, M. R., Canales-Vázquez, J., Castro-García, M., Villas, M. V., Mansilla Legorburo, F., & Sabater, S. (2018). Radiomics of CT features may be nonreproducible and redundant: Influence of CT acquisition parameters. Radiology, 288(2), 407–415. https://doi.org/10.1148/radiol.2018180061

Berger, T. (1971). Rate distortion theory: A mathematical basis for data compression. Prentice-Hall.

Blau, Y., & Michaeli, T. (2018). The perception-distortion tradeoff. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (pp. 6228–6237). https://doi.org/10.1109/CVPR.2018.00652

Brooks, R. A., & Di Chiro, G. (1976). Beam hardening in x-ray reconstructive tomography. Physics in Medicine & Biology, 21(3), 390–398. https://doi.org/10.1088/0031-9155/21/3/004

Bushberg, J. T., Seibert, J. A., Leidholdt, E. M., Jr., & Boone, J. M. (2011). The essential physics of medical imaging (3rd ed.). Lippincott Williams & Wilkins.

Carlsson, G. (2009). Topology and data. Bulletin of the American Mathematical Society, 46(2), 255–308. https://doi.org/10.1090/S0273-0979-09-01249-X

Clark, K., Vendt, B., Smith, K., Freymann, J., Kirby, J., Koppel, P., Moore, S., Phillips, S., Maffitt, D., Pringle, M., Tarbox, L., & Prior, F. (2013). The Cancer Imaging Archive (TCIA): Maintaining and operating a public

information repository. Journal of Digital Imaging, 26(6), 1045–1057. https://doi.org/10.1007/s10278-013-9622-7

Cosman, P. C., Gray, R. M., & Olshen, R. A. (2000). Evaluating quality of compressed medical images: SNR, subjective rating, and diagnostic accuracy. Proceedings of the IEEE, 88(6), 1765–1787. https://doi.org/10.1109/5.900257

Cover, T. M., & Thomas, J. A. (2006). Elements of information theory (2nd ed.). Wiley. https://doi.org/10.1002/047174882X

Deans, S. R. (2007). The Radon transform and some of its applications. Dover Publications. (Original work published 1983)

DeGrave, A. J., Janizek, J. D., & Lee, S.-I. (2021). AI for radiographic COVID-19 detection selects shortcuts over signal. Nature Machine Intelligence, 3(7), 610–619. https://doi.org/10.1038/s42256-021-00338-z

Edelsbrunner, H., & Harer, J. L. (2010). Computational topology: An introduction. American Mathematical Society.

Eisenhauer, E. A., Therasse, P., Bogaerts, J., Schwartz, L. H., Sargent, D., Ford, R., Dancey, J., Arbuck, S., Gwyther, S., Mooney, M., Rubinstein, L., Shankar, L., Dodd, L., Kaplan, R., Lacombe, D., & Verweij, J. (2009). New response evaluation criteria in solid tumors: Revised RECIST guideline (version 1.1). European Journal of Cancer, 45(2), 228–247. https://doi.org/10.1016/j.ejca.2008.10.026

Elbakri, I. A., & Fessler, J. A. (2002). Statistical image reconstruction for polyenergetic X-ray computed tomography. IEEE Transactions on Medical Imaging, 21(2), 89–99. https://doi.org/10.1109/42.993135

Feeman, T. G. (2010). The mathematics of medical imaging. Springer New York.

Fessler, J. A. (2000). Statistical image reconstruction methods for transmission tomography. In M. Sonka & J. M. Fitzpatrick (Eds.), Handbook of medical imaging, Volume 2: Medical image processing and analysis (pp. 1–70). SPIE Press. https://doi.org/10.1117/3.831079.ch1

Gersho, A., & Gray, R. M. (1992). Vector quantization and signal compression. Kluwer Academic.

Gillies, R. J., Kinahan, P. E., & Hricak, H. (2016). Radiomics: Images are more than pictures, they are data. Radiology, 278(2), 563–577. https://doi.org/10.1148/radiol.2015151169

Gökbay, İ. Z. (2013). Modeling biomedical images with the SYMPES method [Doctoral dissertation, Istanbul University].

Goyal, V. K. (2001). Theoretical foundations of transform coding. IEEE Signal Processing Magazine, 18(5), 9–21. https://doi.org/10.1109/79.952802

Hadamard, J. (1902). Sur les problèmes aux dérivées partielles et leur signification physique. Princeton University Bulletin, 13(4), 49–52.

Hadamard, J. (1923). Lectures on Cauchy's problem in linear partial differential equations. Yale University Press.

Hansen, P. C. (1998). Rank-deficient and discrete ill-posed problems: Numerical aspects of linear inversion. Society for Industrial and Applied Mathematics. https://doi.org/10.1137/1.9780898719697

Hosny, A., Parmar, C., Quackenbush, J., Schwartz, L. H., & Aerts, H. J. W. L. (2018). Artificial intelligence in radiology. Nature Reviews Cancer, 18(8), 500–510. https://doi.org/10.1038/s41568-018-0016-5

Hsieh, J. (2009). Computed tomography: Principles, design, artifacts, and recent advances (2nd ed.). SPIE Press. https://doi.org/10.1117/3.832732

Ilyas, A., Santurkar, S., Tsipras, D., Engstrom, L., Tran, B., & Madry, A. (2019). Adversarial examples are not bugs, they are features. In Advances in Neural Information Processing Systems 32 (pp. 125–136). https://doi.org/10.48550/arXiv.1905.02175

Kak, A. C., & Slaney, M. (2001). Principles of computerized tomographic imaging. Society for Industrial and Applied Mathematics. https://doi.org/10.1137/1.9780898719277

Keall, P. J., Mageras, G. S., Balter, J. M., Emery, R. S., Forster, K. M., Jiang, S. B., Kapatoes, J. M., Low, D. A., Murphy, M. J., Murray, B. R., Ramsey, C. R., Van Herk, M. B., Vedam, S. S., Wong, J. W., & Yorke, E. (2006). The management of respiratory motion in radiation oncology report of AAPM Task Group 76. Medical Physics, 33(10), 3874–3900. https://doi.org/10.1118/1.2349696

Kramme, R., Hoffmann, K. P., & Pozos, R. S. (Eds.). (2011). Springer handbook of medical technology. Springer Science & Business Media.

Lambin, P., Leijenaar, R. T., Deist, T. M., Peerlings, J., De Jong, E. E., Van Timmeren, J., Sanduleanu, S., Larue, R. T. H. M., Even, A. J. G., Jochems, A., van Wijk, Y., Woodruff, H., van Soest, J., Lustberg, T., Roelofs, E., van Elmpt, W., Dekker, A., Mottaghy, F. M., Wildberger, J. E., & Walsh, S. (2017). Radiomics: The bridge between medical imaging and personalized medicine. Nature Reviews Clinical Oncology, 14(12), 749–762. https://doi.org/10.1038/nrclinonc.2017.141

Lambin, P., Rios-Velazquez, E., Leijenaar, R., Carvalho, S., van Stiphout, R. G., Granton, P., Zegers, C. M. L., Gillies, R., Boellard, R., Dekker, A., & Aerts, H. J. W. L. (2012). Radiomics: Extracting more information from medical images using advanced feature analysis. European Journal of Cancer, 48(4), 441–446. https://doi.org/10.1016/j.ejca.2011.11.036

Mackin, D., Fave, X., Zhang, L., Fried, D., Yang, J., Taylor, B., Rodriguez-Rivera, E., Dodge, C., Jones, A. K., & Court, L. (2015). Measuring

computed tomography scanner variability of radiomics features. Investigative Radiology, 50(11), 757–765. https://doi.org/10.1097/RLI.0000000000000180

Macovski, A. (1983). Medical imaging systems. Prentice-Hall.

Natterer, F. (2001). The mathematics of computerized tomography. Society for Industrial and Applied Mathematics. https://doi.org/10.1137/1.9780898719239

Radon, J. (1917). Über die Bestimmung von Funktionen durch ihre Integralwerte längs gewisser Mannigfaltigkeiten. Berichte über die Verhandlungen der Königlich-Sächsischen Akademie der Wissenschaften zu Leipzig, Mathematisch-Physische Klasse, 69, 262–277.

Rudin, L. I., Osher, S., & Fatemi, E. (1992). Nonlinear total variation based noise removal algorithms. Physica D: Nonlinear Phenomena, 60(1–4), 259–268. https://doi.org/10.1016/0167-2789(92)90242-F

Sauer, K., & Bouman, C. (1993). A local update strategy for iterative reconstruction from projections. IEEE Transactions on Signal Processing, 41(2), 534–548. https://doi.org/10.1109/78.193196

Sayood, K. (2017). Introduction to data compression (5th ed.). Morgan Kaufmann.

Shannon, C. E. (1948). A mathematical theory of communication. Bell System Technical Journal, 27(3), 379–423. https://doi.org/10.1002/j.1538-7305.1948.tb01338.x

Sheikh, H. R., & Bovik, A. C. (2006). Image information and visual quality. IEEE Transactions on Image Processing, 15(2), 430–444. https://doi.org/10.1109/TIP.2005.859378

Sidky, E. Y., & Pan, X. (2008). Image reconstruction in circular cone-beam computed tomography by constrained, total-variation minimization. Physics in Medicine & Biology, 53(17), 4777–4807. https://doi.org/10.1088/0031-9155/53/17/021

Tikhonov, A. N., & Arsenin, V. Y. (1977). Solutions of ill-posed problems. V. H. Winston & Sons.

Tishby, N., Pereira, F. C., & Bialek, W. (2000). The information bottleneck method. In Proceedings of the 37th Annual Allerton Conference on Communication, Control, and Computing (pp. 368–377). University of Illinois.

Tryggestad, E. J., Liu, W., Pepin, M. D., Hallemeier, C. L., & Sio, T. T. (2020). Managing treatment-related uncertainties in proton beam radiotherapy for gastrointestinal cancers. Journal of Gastrointestinal Oncology, 11(1), 212–224. https://doi.org/10.21037/jgo.2019.11.07

Wang, Z., Bovik, A. C., Sheikh, H. R., & Simoncelli, E. P. (2004). Image quality assessment: From error visibility to structural similarity. IEEE

Transactions on Image Processing, 13(4), 600–612. https://doi.org/10.1109/TIP.2003.819861

Zech, J. R., Badgeley, M. A., Liu, M., Costa, A. B., Titano, J. J., & Oermann, E. K. (2018). Variable generalization performance of a deep learning model to detect pneumonia in chest radiographs: A cross-sectional study. PLOS Medicine, 15(11), e1002683. https://doi.org/10.1371/journal.pmed.1002683

Medical Image Compression in the Era of Clinical AI:

The Fidelity–Utility Dilemma and Information Preservation

Asst. Prof. Dr. İnci Zaim Gökbay
Prof. Dr. Bekir Sıddık Binboğa Yarman

ÖZGÜR PRESS