

# Investigating the Propagation of CT Acquisition Artifacts along the Medical Imaging Pipeline

Jakob Peischl and Renata G. Raidou<sup>a</sup>

*Institute of Visual Computing & Human-Centered Technology, TU Wien, Favoritenstrasse 9-11, 1060, Vienna, Austria*

**Keywords:** Uncertainty Quantification and Propagation, Medical Imaging, Medical Image Analysis and Visualization.

**Abstract:** We propose a framework to support the simulation, exploration, and analysis of uncertainty propagation in the medical imaging pipeline—exemplified with artifacts arising during CT acquisition. Uncertainty in the acquired data can affect multiple subsequent stages of the medical imaging pipeline, as artifacts propagate and accumulate along the latter, influencing the diagnostic power of CT and potentially introducing biases in eventual decision-making processes. We designed and developed an interactive visual analytics framework that simulates real-world CT artifacts using mathematical models, and empowers users to manipulate parameters and observe their effects on segmentation outcomes. By extracting radiomics features from artifact-affected segmented images and analyzing them using dimensionality reduction, we uncover distinct patterns related to individual artifacts or combinations thereof. We demonstrate our proposed framework on use cases simulating the effects of individual and combined artifacts on segmentation outcomes. Our application supports the effective and flexible exploration and analysis of the impact of uncertainties on the outcomes of the medical imaging pipeline. Initial insights into the nature and patterns of the simulated artifacts could also be derived.

## 1 INTRODUCTION

Medical imaging is foundational to modern medicine, aiding in disease prevention, diagnosis, and treatment. Imaging modalities like Computed Tomography (CT) and Magnetic Resonance Imaging (MRI) provide patient-specific information based on physical principles that capture dedicated tissue characteristics to support clinical decisions (Gillmann et al., 2021). However, *data uncertainties* stemming from several factors, such as low spatial resolution, artifacts, or hardware/software issues, are often present (Ristovski et al., 2014; Gillmann et al., 2021). This causes discrepancies between the physical properties of imaged tissue and the respective image representation. These uncertainties are subsequently *propagated* through the medical imaging pipeline, affecting all the remaining steps of processing, segmentation, mapping, and rendering (Preim and Botha, 2013)—potentially, impacting the accuracy of diagnosis and treatment.

Uncertainty in medical imaging can affect multiple stages of the medical imaging pipeline, with significant challenges at each step (Schlachter et al., 2019). For example, in CT data acquisition we often encounter artifacts due to patient motion and partial

volume artifacts. Such artifacts may further propagate and accumulate during pre-processing, segmentation, and mapping, influencing the diagnostic power of the employed data and potentially introducing biases in eventual decision-making processes (Raidou, 2018). For decisions involving artificial intelligence (AI), uncertainty can additionally compromise model reliability (Zhou et al., 2021). Whether the “decision maker” is an AI model or a human, *recognizing and mitigating* uncertainties is essential.

This paper focuses on *understanding* the propagation of acquisition uncertainty, exemplified with CT imaging. CT is integral to numerous clinical processes, for which the presence of potential uncertainties is often critical. Motion artifacts, partial volume effects, and ring artifacts are among the most prevalent uncertainties that impact CT imaging quality (Preim and Botha, 2013). By examining these artifacts, we aim to map the ways uncertainties in the acquisition step of the medical imaging pipeline affect subsequent steps—and in particular, the last step of analysis.

To this end, we introduce a *methodology* to simulate and investigate the propagation of CT acquisition uncertainty using mathematical models that describe the aforementioned, common artifacts. Our proposed

<sup>a</sup>  <https://orcid.org/0000-0003-2468-0664>

methodology involves a user-in-the-loop approach for constructing uncertainty pipelines by combining artifacts to model their effects on the acquired images, and their impact on the subsequent segmentation. Making use of radiomics feature extraction and unsupervised learning approaches, we visualize and analyze segmentation outcomes from artifact-affected images. In this way, we can evaluate in real time how acquisition uncertainty has propagated through the pipeline and the magnitude of its impact on diagnostic information. Although our focus is on CT artifacts, the proposed methodology can be adapted to other modalities, allowing for broader applicability in medical imaging uncertainty research.

Our *contributions* to the field of uncertainty visualization include the introduction of a novel framework for acquisition uncertainty analysis in medical imaging, an interactive application for the investigation of simulated uncertainty pipelines, and a demonstration of use cases that illustrate the potential of our approach in a real-life setting.

## 2 RELATED WORK

In this section, we provide an overview of different uncertainty types and examine relevant literature on uncertainty quantification and propagation, with a focus on medical imaging.

**Uncertainty types** are classified as either *epistemic* or *aleatoric* (Potter et al., 2012). Epistemic uncertainty results from model imperfections and can be reduced by improved measurements or calibration, while aleatoric uncertainty arises from inherent randomness and is typically represented with probability density functions (PDFs). Given that CT image artifacts are often reducible or avoidable, our work focuses on epistemic uncertainty. However, our employed methods could theoretically extend to aleatoric uncertainties.

We target specific **medical imaging pipeline steps** that we anticipate to be affected by uncertainties (Gillmann et al., 2021). Each step introduces new uncertainties: acquisition uncertainties may arise from physical model assumptions, segmentation algorithms add model-based and parameter uncertainties affecting boundary delineations, and mapping, e.g., via marching cubes (Lorensen and Cline, 1987), introduces approximations to the derived surface meshes. In this study, segmentation and mapping are kept invariant to investigate in an isolated manner uncertainties only from the acquisition step.

(Ristovski et al., 2014) and (Gillmann et al., 2021) recently authored surveys for **uncertainty classifica-**

**tion in medical applications.** Ristovski et al. classify uncertainties using random fields to represent arbitrary probability distribution functions. By determining the mathematical characteristics of a random field, they provide a framework to consider uncertainty propagation behavior. Gillmann et al. review state-of-the-art uncertainty-aware visualization techniques related to medical imaging applications, explore pipeline combinations, and identify challenges in creating comprehensive uncertainty-aware medical imaging workflows.

**Uncertainty quantification and propagation** requires formalized mathematical models to track how uncertainties propagate within an imaging pipeline. A general formulation utilizes PDFs, where each input  $X_i$  of the input vector  $\mathbf{X} = (X_1, \dots, X_n)^T$  is given as a PDF  $f_{X_i}(x_i)$ . The challenge is then to find the PDF of the output  $\mathbf{Y}$  after it has been transformed by the function  $\mathbf{g}$ , i.e., finding  $\mathbf{f}_Y(\mathbf{y})$  where  $\mathbf{Y} = \mathbf{g}(\mathbf{X})$ . Various approaches for determining the output uncertainty in the so-called *forward uncertainty quantification* problem exist (Brodlić et al., 2012). For complex uncertainties, Monte-Carlo-based techniques that approximate  $\mathbf{f}_Y(\mathbf{y})$  are often necessary (Brodlić et al., 2012). A primary drawback of these methods is computational inefficiency due to slow convergence. Analytical solutions would be more efficient but are unavailable for arbitrary PDF propagation.

Uncertainty quantification has become a relevant aspect of a wide range of scientific fields (Zhang, 2021). **Uncertainty-handling frameworks** have been proposed for instance by (Roy and Oberkampf, 2011) to model, estimate, and propagate through any scientific model uncertainties. (Wu et al., 2012) focus on explorative aspects of uncertainty and build an interactive tool that visualizes uncertainty flows along a data processing pipeline—under the assumption that input uncertainties follow a multivariate normal distribution. Again, the output uncertainty is estimated using Monte Carlo sampling.

The application of uncertainty quantification within medical imaging has primarily focused on **visualization and single-stage analysis** of uncertainty within images. For example, (Howard et al., 2014) use sampling-based CT uncertainty quantification, and (Gillmann et al., 2017) employ analytical methods to highlight image noise as a primary uncertainty source. (Tian and Samei, 2016) model CT quantum noise, and (Gravel et al., 2004) examine noise profiles across modalities. (Diwakar and Kumar, 2018) consolidate insights into CT noise, highlighting denoising as a key challenge.

There is limited literature on **uncertainty propagation** across multiple stages of medical imaging

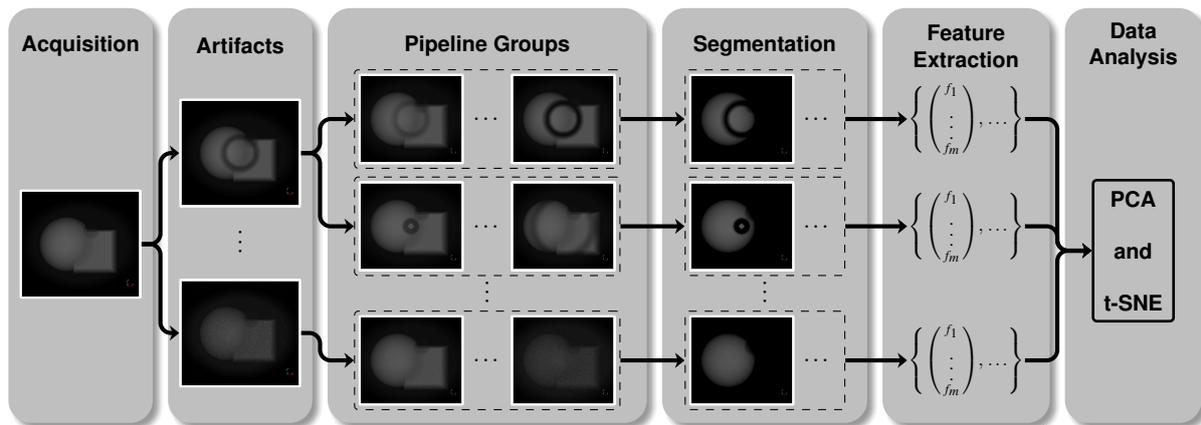


Figure 1: Overview of the workflow adopted in our approach. The workflow comprises five steps: *acquisition*, *artifact generation*, *pipeline grouping*, *segmentation*, and *outcome analysis* (which consists of two substeps: *radiomics feature extraction* and *analysis* based on dimensionality reduction).

pipelines. A notable exception is Diffusion Tensor Imaging (DTI), where (Behrens et al., 2003) laid the groundwork for uncertainty propagation from diffusion-weighted signals to tensor-derived quantities. Yet, DTI findings are not generalizable to CT due to modality-specific differences in imaging reconstruction and clinical use cases. Recent advancements in deep learning (DL) have created interest in both epistemic and aleatoric uncertainty quantification and propagation within medical imaging (Feiner et al., 2023). However, CT artifacts affecting DL models in real-world scenarios remain underexplored. Only (Athanasίου et al., 2015) explored multi-stage error propagation within imaging pipelines for plaque classification, linking acquisition and segmentation errors to classification accuracy. In contrast, our work evaluates outcomes based on radiomics feature comparisons to obtain a holistic view of CT uncertainties across pipeline stages.

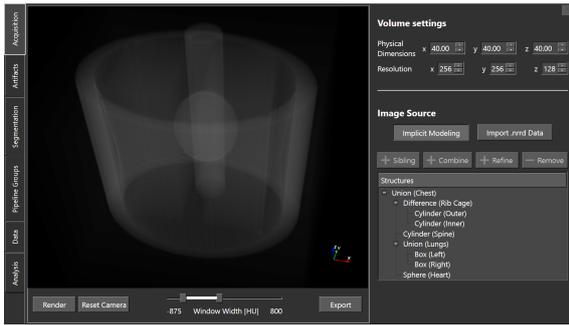
### 3 METHODOLOGY

In this chapter, we present our proposed methodology for the interactive quantification and propagation of CT uncertainties through the medical imaging pipeline. The workflow involves five stages, which are depicted in Fig. 1 and further described in the upcoming subsections. We first generate an uncertainty-free base image during the *image acquisition* stage (Sec. 3.1). Users can then add artifacts, whose effects are applied to this base image in the *artifacts generation* stage (Sec. 3.2). The *pipeline grouping* stage allows users to combine different artifacts and specify valid parameters for each of them (Sec. 3.3). We then proceed with the *segmentation* step to calculate seg-

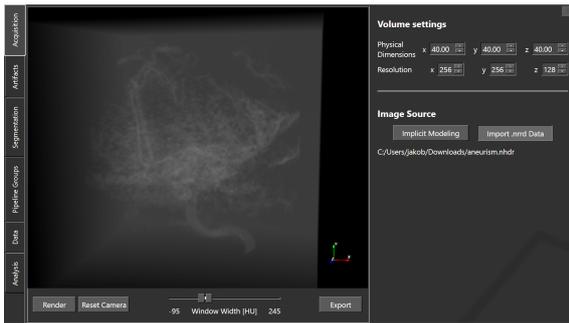
mentations of the artifact-affected images (Sec. 3.4). Next, we *extract radiomics features* from the outcome segmentations and comparatively *analyze* them using a combination of principal component analysis (PCA) and t-distributed stochastic neighbor embedding (t-SNE), inspired by the previous work of (Reiter et al., 2018) (Sec. 3.5). This is the *outcome analysis* stage. Our interface facilitates these steps, targeting medical imaging researchers and CT imaging professionals interested in uncertainty quantification and propagation.

#### 3.1 Image Acquisition

At the image acquisition stage, we define the extent, i.e. the physical dimensions, and spatial resolution of the CT input volume. A CT dataset can be defined mathematically as a function  $I: \mathbb{R}^3 \rightarrow \mathbb{R}$  that maps a three-dimensional spatial position  $(x, y, z)$  to a corresponding voxel value  $I(x, y, z)$ . CT voxel values (i.e. radiodensity values) are given in Hounsfield Units (HU), proportional to the attenuation coefficient normalized to water (Yucel-Finn et al., 2023). Users can define voxel values through *implicit modeling* or *direct image import*. In the former case, users employ constructive solid geometry to model complex objects from primitive shapes (Laidlaw et al., 1986). This approach provides structural information for subsequent stages and it is showcased in Fig. 2 (a). In the latter case, users can import image data allowing for a more realistic data analysis (Fig. 2 (b)). When choosing this option, artifact simulation might be limited, and the image may already suffer from some kind of uncertainty. For simplicity, our workflow assumes an idealized, artifact-free CT acquisition.



(a)



(b)

Figure 2: In the *image acquisition* stage, the user can generate data by (a) *implicit modeling* or by (b) *direct image import* (here, a CT angiogram containing a brain aneurysm (Klacansky, 2017)).

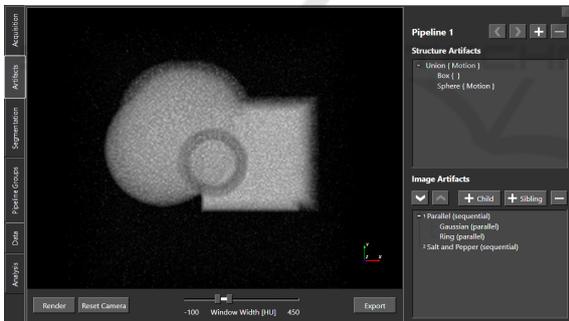


Figure 3: In the *artifacts generation* stage, the user defines which artifacts are to be simulated and in which order. Here, two artifacts are simulated and combined: a *motion artifact* (comprising a Gaussian blur and a ring artifact), and *salt and pepper noise*. The order in which artifacts are applied is specified in the bottom-right corner of the interface.

## 3.2 Artifacts Generation

In this stage, users define which artifacts they would like to simulate on the data obtained from the first step (Fig. 3). Rather than simulating CT acquisition, we assume an artifact-free CT image, to which artifact effects are added. This ensures a flexible simulation across scanning technologies by adjusting artifact parameters. Multiple artifact sets can also be created

and are combined later, in the stage where pipeline groups are simulated (Sec. 3.3).

Our framework separates potential CT artifacts into two types: *image artifacts* and *structure artifacts* (Fig. 4). Image artifacts apply to the entire CT volume based on voxel-specific radiodensity without needing structural information. For example, a Gaussian noise artifact adds random noise to each voxel, as determined by a user-defined Gaussian distribution. Conversely, structure artifacts affect specific structures (or subregions within an image); thus, they need additional positional information. For instance, a motion artifact would affect a structure by shifting it and by applying a Gaussian blur, as shown in Fig. 3. Simulating this is only possible if the positional information of the affected structure is known. Both types of artifacts are further discussed below.

Artifacts can be applied *sequentially* or *in parallel* on a given dataset. For image artifacts, we can apply artifacts sequentially, in parallel, or in a (potentially nested) combination of both approaches, as shown in the bottom-right corner of the interface in Fig. 3. In contrast, the effects of structure artifacts are applied only in parallel because they are inherently localized to specific structures within the CT image. Applying them sequentially would not make sense because each structure artifact alters independently its associated structure without influencing others.

### 3.2.1 Image Artifacts

The image artifacts modeled within our approach, together with their parameters, are summarized in Fig. 4 (a)–(d). Each of these artifacts follows mathematical descriptions, as formalized in literature and included in our *Appendix*. **Salt and pepper noise** appears as isolated bright (“salt”) or dark (“pepper”) voxels, caused by rare electronic errors and is generally minor in CT imaging (Lu et al., 2018) (Fig. 4 (a)). Despite its limited impact, we include it here for completion. Parameters of our modeling include *radiodensity* and the *relative amount* of salt and pepper pixels. **Gaussian noise**, observed in low-dose CT images, results from low photon counts and other electronics-based noise (Boas and Fleischmann, 2012) (Fig. 4 (b)). Though often modeled by a Poisson distribution, a Gaussian distribution suffices at high photon counts (Madhura and Babu, 2017). The parameters of our modeling are the *mean* and *standard deviation* of the distribution. **Cupping artifacts**, due to beam hardening, appear as shading towards the center of a dense object (e.g. a hard bone structure) (Barrett and Keat, 2004) (Fig. 4 (c)). The shading appears pronounced in images with larger, isotropic objects like the skull. Its modeling relies on user-defined param-

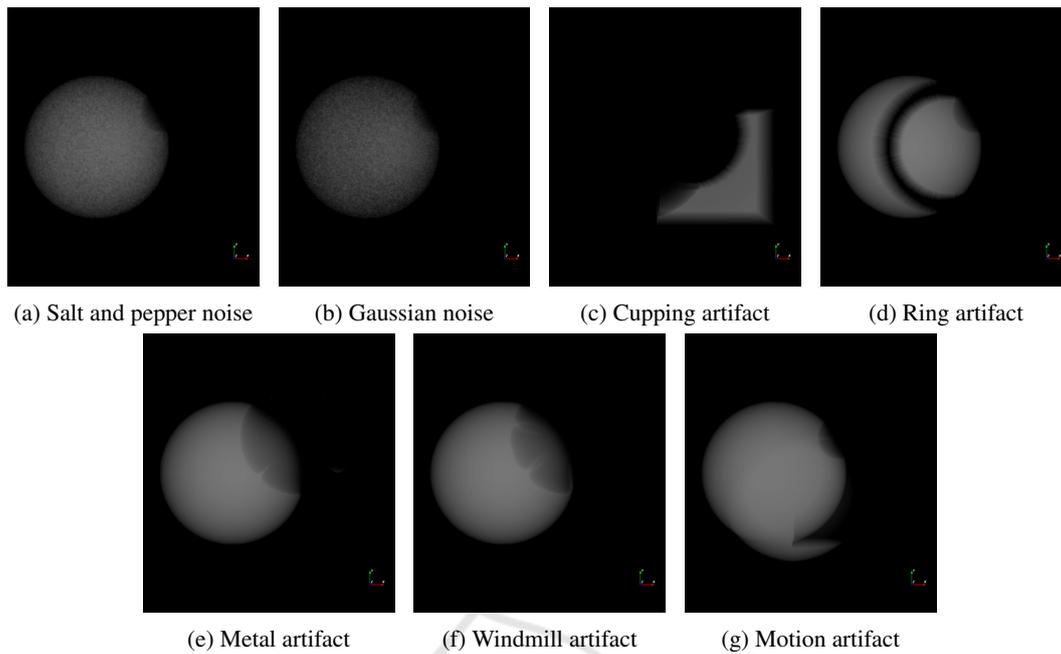


Figure 4: The artifacts that can be generated as part of our approach, exemplified on a simple implicitly modeled data set. Subfigures (a)–(d) represent *image artifacts*, while (e)–(g) represent *structure artifacts*.

eters for *center location* and *shading intensity*. Finally, **ring artifacts** are caused by detector miscalibrations, forming rings around the scanning axis, and are sometimes mistaken for pathologies (Boas and Fleischmann, 2012) (Fig. 4 (d)). Modeling parameters include the *inner radius*, *width*, *radiodensity*, and *ring center location*.

### 3.2.2 Structure Artifacts

The structure artifacts modeled within our approach, together with their parameters, are summarized in Fig. 4 (e)–(g). Each of these artifacts follows mathematical descriptions, as formalized in literature and included in our *Appendix*. **Metal artifacts** are common in CT due to metal objects that cannot be removed during acquisition (Boas and Fleischmann, 2012) (Fig. 4 (e)). These artifacts stem from effects such as beam hardening, which causes directional shading and photon scatter, making shading more prominent. This artifact is modeled with assumptions on attenuation and distance effects, using parameters for *attenuation direction* and *shading length*. **Windmill artifacts** appear as evenly spaced, bright streaks around high-attenuation objects (Barrett and Keat, 2004) (Fig. 4 (f)). They occur only in helical CT and increase in number with greater pitch. The model assumes shading decreases with distance, using parameters for *radiodensity*, *length*, and *rotation per slice*. Finally, **motion artifacts** are frequent, of-

ten due to involuntary or internal movement of the patient, showing as blurring or double edges (Boas and Fleischmann, 2012) (Fig. 4 (g)). The model captures these by transforming structure locations and applying Gaussian blur to regions of motion, using parameters for *transformation matrix* and *blur properties*.

### 3.3 Pipeline Grouping

At this stage, final artifact parameters are set for customizable pipeline groups. Each group contains combinations of artifact pipelines (serial, in parallel, or (nested) combinations thereof) based on a base pipeline, which users adjust with parameter ranges. The primary purpose of the parameter ranges is to determine a set of valid values for an artifact parameter. For example, this enables a user to vary the standard deviation of a Gaussian artifact contained in the base pipeline in an interval  $[a, b]$  and a step size  $s$ , resulting in  $n = \lfloor \frac{b-a}{s} \rfloor$  distinct artifact pipelines, as shown in Fig. 1. Since multiple parameter ranges may be defined per pipeline group and artifact, the total number of valid configurations for a pipeline group corresponds to the number of elements in the Cartesian product of the parameter spans.

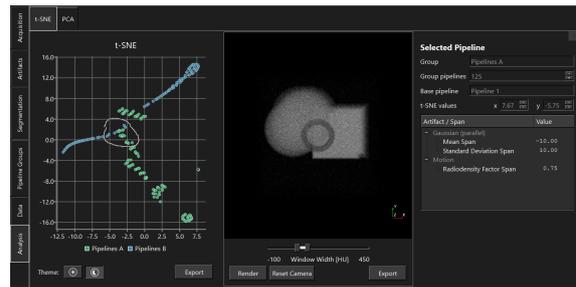
### 3.4 Segmentation

During the segmentation stage, the user can segment a structure of interest. As we are particularly interested in the propagation of acquisition uncertainties through the pipeline and their effect on the outcomes, we choose to keep the segmentation filter simple and its parameters unchanged. Therefore, we employ a simple thresholding approach where the user only needs to select an appropriate threshold for the segmentation. An example of the segmentation outcome of different pipeline groupings is shown in Fig. 1. This ensures that we can later compare the resulting feature values meaningfully since all images have undergone the same segmentation—without having to additionally consider the effects of the segmentation step itself on the uncertainty outcomes.

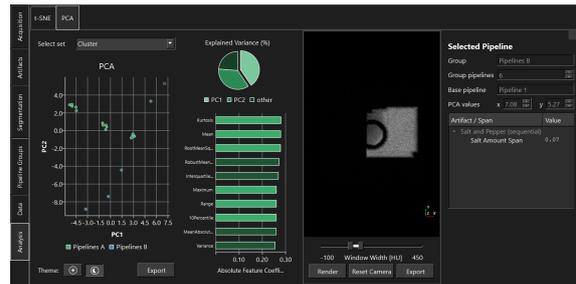
### 3.5 Outcome Analysis

A quantitative comparison of the artifact effects on the outcome of the pipeline is not feasible using segmented images alone. To this end, we adopt a *radiomics-based approach*, extracting predefined radiomics features from the segmented outcome of each artifact-impacted pipeline. Radiomics are quantitative metrics derived from medical images and are often used to characterize tumors (Gillies et al., 2016). They consist of a series of features that describe the morphology, texture, and intensity of a segmentation. We extract 13 3D shape features and 19 first-order features, chosen for their time complexity of  $O(n \log n)$  and their simpler physical interpretations.

Subsequently, to *comparatively explore and analyze* relationships between artifact pipelines, we employ a t-distributed stochastic neighborhood embedding (t-SNE) (van der Maaten and Hinton, 2008) and principal component analysis (PCA) (Abdi and Williams, 2010), similar to the prior work of (Reiter et al., 2018). Using a Barnes-Hut algorithm for t-SNE (van der Maaten, 2014) allows for a reduced time complexity of  $O(n \log n)$ . We standardize the feature data for equal weighting and apply t-SNE to the entire dataset *across* pipeline groups (e.g. multiple artifacts with multiple parameterizations) while using PCA *within* pipeline groups (e.g. one type of artifact with several parameterizations), as shown in Fig. 5. The t-SNE results are represented in an interactive 2D scatterplot, with data points color-coded by pipeline group (Fig. 5 (a)). Selecting a subset of points in the t-SNE plot triggers a PCA analysis on this filtered selection. The PCA output is also shown in a 2D scatterplot, with insights into the PCs and their significance using a pie and bar chart (Fig. 5 (b)). This



(a)



(b)

Figure 5: In the *outcome analysis* step, (a) we use a *t-SNE analysis* to compare pipeline groups. The left window displays a scatterplot of t-SNE results for radiomic features from each artifact-impacted pipeline. Selecting a data point in this plot reveals the associated pipeline in the right window and renders the affected data in the center window. Here, a cluster of points is selected on the t-SNE view using a lasso brush. (b) Next, we perform *PCA analysis* on the lasso-selected points in (a). The left window shows a scatterplot of the first two PCs. The center and right windows present a pie chart of explained variances for these PCs with a bar chart of sorted feature coefficients for PC1, and renderings and pipeline details for the selected data. Dark green indicates negative feature coefficients.

combination leverages the ability of t-SNE to detect non-linear relationships across pipelines and the interpretability of PCA within pipelines.

### 3.6 Implementation

For the implementation of the framework, C++20 was chosen for its performance capabilities and low-level control, while the Qt framework is used for developing a cross-platform graphical user interface. For medical image processing, the application employs the Visualization Toolkit (VTK). Data storage utilizes the HDF5 format with zlib compression to manage the large memory requirements effectively. For feature extraction, PyRadiomics is used. The implementation of our approach is available on our Open-Source Repository.

## 4 RESULTS

This section presents the results of our uncertainty propagation framework applied to two scenarios. We describe each scenario and analyze the results, followed by a cumulative summary and discussion. The selected scenarios are intended to demonstrate the capabilities of our proposed solution with examples representing both implicitly modeled and imported data. Each scenario was designed to enable insights relevant to real-world cases, using CT data and artifact parameters that reflect realistic parameter ranges.

### 4.1 Scenario 1: Modeled Data

The first example scenario uses *implicitly modeled CT data*. We created a CT structure scene that is challenging for segmentation by placing structures with similar radiodensities close to a structure of interest. We added both *image and structure artifacts* (see Fig. 4), making segmentation more complex and analysis more meaningful for a real-life scenario. This setup aims to simulate applications like cardiac and pelvic organ CT segmentation, where similar radiodensities and metal artifacts (e.g., from pacemakers or hip replacements) commonly introduce uncertainties complicating processing and analysis tasks. *Segmentation* was performed using a bandpass threshold filter with a specific radiodensity range. Although more complex segmentation algorithms could be used, we consider this out of the scope of this work.

For this first example scenario, we analyzed the effects of individual artifacts in isolation. Each *pipeline group* contains only artifact pipelines with the same type of artifact. This way, we can explore all artifacts simultaneously while keeping the complexity of the example manageable. For each group, artifacts like metal and windmill distortions (see Fig. 4 (e–f)) have been applied to relevant structures, while a motion artifact has been introduced to affect the appearance of the sphere of interest. By *varying specific artifact parameters* within literature-defined ranges, we generate seven pipeline groups—each targeting a different artifact. After parameterization, we have a total of 752 individual pipelines to explore and analyze.

We then proceed with extracting the 32 *radiomics features* for each case (13 3D shape features and 19 first-order features). Running a *Barnes-Hut-based t-SNE* algorithm (van der Maaten, 2014) on the standardized feature data extracted from the artifact-affected images results in a clear separation of the pipeline groups and, therefore, the different artifact types (Fig. 6). Nearly all data points of individual pipeline groups lie within clusters near each other

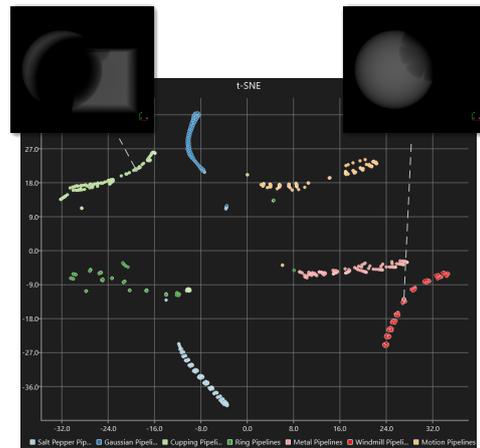


Figure 6: t-SNE for Scenario 1. Small relative distances suggest similarities between data points. The segmentation renderings of two pipelines are provided as examples.

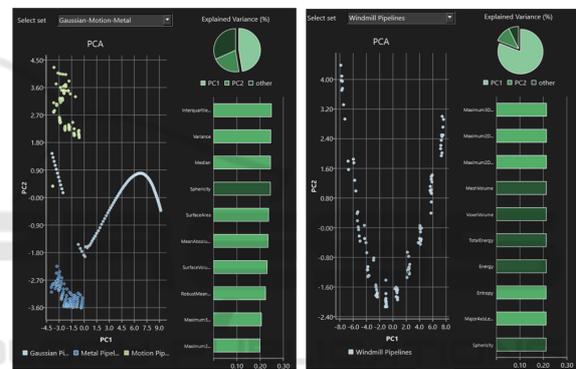


Figure 7: Inter-pipeline-group PCAs for Scenario 1: (left) between Gaussian noise, motion, and metal pipelines; (right) for windmill artifact pipelines.

while being at a distance from data points corresponding to another artifact type. The outliers result from extreme parameter values in the artifact generation; thus, are not discussed any further. Another interesting observation is that for all pipeline group clusters, the data points are stretched along one main direction, with one end of the cluster corresponding to the largest segmentations within the group and the other end to the smallest ones. Additionally, for the pipeline groups for which two artifact parameters were varied (all except for the Gaussian noise), we also noticed the presence of sub-clusters—potentially indicating competing interests between parameter setups.

Further *analysis with PCA* offers insights into how specific artifact parameters impact segmentation (Fig. 7). We take a closer look at the PCA for the windmill pipelines only (Fig. 7, right), as an example for deriving insights about the impact of the two windmill artifact characteristics (i.e. length and angu-

lar width) on the segmentation. PCA shows the same sub-clustering pattern observed for the t-SNE, indicating that the method can separate variations of the two artifact parameters along two dimensions. Analyzing the individual data points, we see that the length increases in the direction of PC1, and the angular width decreases along the direction of PC2.

We examine PC1 more closely since it explains nearly 81% of the variance in the windmill pipelines data (see Fig. 7, right; pie chart) and notice that it assigns equally high priority to many features (see Fig. 7, right; bar chart). Most (21 out of 32) features have an absolute coefficient value within a range [0.18, 0.21]. This indicates a redundancy between the employed features. This observation agrees with expectations of the relationships between individual features (Gillies et al., 2016). For instance, the total energy depends on the energy by a factor of the voxel volume.

Overall, the shape features dominate the first-order features (Fig. 7, bar charts left and right). PCA achieves a precise separation of pipelines by looking at individual features. With increasing length, the windmill artifact affects larger portions of the sphere of interest, leading to smaller segmentations. Also, an increasing length widens the gray-level histogram of the image due to the gradual nature of the windmill artifact, which explains the high positive weighting for the radiomics feature of entropy. On the other hand, quantile-related features provide very little information for PC1 with median or interquartile range having coefficients close to 0. This is probably due to the substantial uniformity in the image, given the assumption that the sphere of interest was modeled as a homogeneous region.

**Performance.** Generating the artifact-affected CT images took the largest time portion (approx. 1 hour 38 minutes). The feature extraction demanded around 16 minutes. The PCA and the t-SNE analysis only took a negligible amount of time compared to the other two tasks. Based on the volume settings and implementation details, the file size of each CT image was determined to be 50.4 MB. For all artifact pipelines, this amounts to 38 GB, which was reduced to less than 3.9 GB by compression.

## 4.2 Scenario 2: Real CT Data

In this second scenario, we employ *real-world CT data* from a head angiography (Klacansky, 2017), specifically focusing on the arteries in the right half of a patient's head, where an aneurysm is present. The CT scan, enhanced with a contrast agent, was chosen due to its relevance in clinical imaging and segmen-

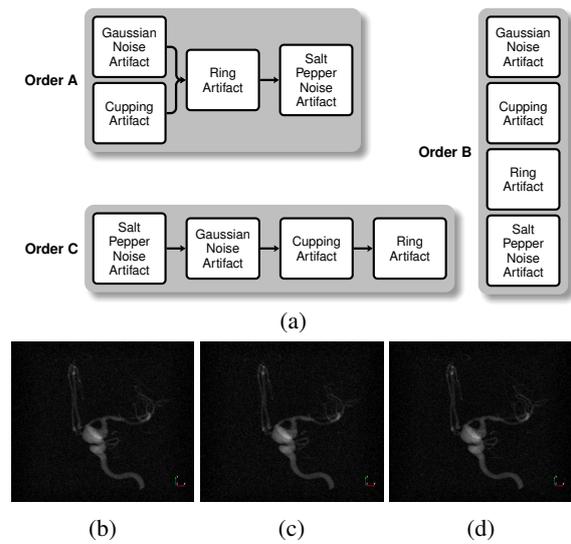


Figure 8: (a) Schematic depiction of the order by which the different artifacts are applied to the dataset in Scenario 2. (b–d) Visual comparison of order A, B, and C, respectively.

tation tasks, particularly in assessing aneurysm shape and volume. The detailed and complex arterial structures make segmentation sensitive to artifacts, and with only one primary structure of interest, the setup remains interpretable. This scenario aims to explore the *impact of artifact application order* on segmentation outcomes, using four types of image artifacts (salt and pepper noise, Gaussian noise, cupping, and ring artifacts). We tested *three pipeline configurations* (Fig. 8 (a)):

**Order A.** This order replicates the order in which artifacts naturally occur in real CT imaging, with Gaussian noise and cupping first (as transmission artifacts), followed by ring (being scanner-based artifacts) and salt-and-pepper noise (again due to data transmission) (see Fig. 8 (b)).

**Order B.** All artifacts are applied in parallel (see Fig. 8 (c)).

**Order C.** All artifacts are applied in sequence, with an arbitrary ordering that is distinct from the other two for diversity (see Fig. 8 (d)).

In each configuration, *artifact parameters* remained consistent, though the sequence varied. Visual analysis of these orders showed slight noise patterns and segmentation differences in fine structures. To determine an optimal threshold (109 HU) for *segmentation*, Otsu's method (Otsu, 1979) was used to maximize the inter-class variance between the background and the structure of interest. The same 32 *radiomics features* were also extracted here for all cases.

The *t-SNE analysis* displayed five relatively distinct clusters, although less well-formed than in the

previous scenario (Fig. 9). The algorithm was able to find similarities within five groups of data points, albeit they were not as strong as for the first scenario. Each cluster included data points from multiple pipeline groups, which means that some inter-pipeline-group similarities exist in our data. This is an expected result because the pipeline groups use identical artifact parameters. However, we also see differences caused by the different orders since Clusters 1, 4, and 5 (Fig. 9) contain unequal shares of data points from each pipeline group. While Clusters 2 and 3 *only* contain data points from Order B and C, respectively. Salt noise was found consistently across intra-group points, while other artifact parameters varied across their full range. Additionally, we find that the position of a data point within a cluster no longer reliably predicts the values of individual artifact parameters. This suggests that clustering patterns emerged from combinations of artifact parameters rather than individual ones.

PCA was used to interpret the parameter impacts on segmentation (Fig. 10). PC1 explained over 80% of the variance in individual groups, primarily due to the absence of salt noise (Fig. 10, *left*; *pie chart*). We already determined this factor to be influential during the t-SNE analysis. PCA hardly reveals any information about the influence of artifact order because the separation affects all pipeline groups approximately the same. Based on the features with the biggest feature coefficients, we observe that many of our features are sensitive to large outliers within the image. In PC2, which explains only 6.7% of the variance (Fig. 10, *left*; *pie chart*) but is less correlated with PC1, entropy and uniformity were primary contributors among the radiomics features. Order C showed the highest entropy, especially in images with salt noise, while Order B exhibited the lowest. Order B, with parallel application, impacts radiodensities less, and Order A places salt and pepper noise last, reducing entropy.

Upon looking at the PCA of pipelines with salt noise (Fig. 10, *right*), shape features contributed significantly to PC1 (Fig. 10, *right*; *pie and bar chart*). Least axis length and minor axis length affect PC1 negatively. Here, these can be interpreted as preferring a center of mass that aligns better with the center of the bounding box of a segmentation. This is the case when the ring artifact affects the region of the aneurysm strongly because it is the thickest structure in the image, which can reduce the size of the error ellipsoid. These quantities stay relatively uniform for Order B compared to Orders A and C because with the parallel artifacts of order B, the effects of one artifact do not influence the effects of the others.

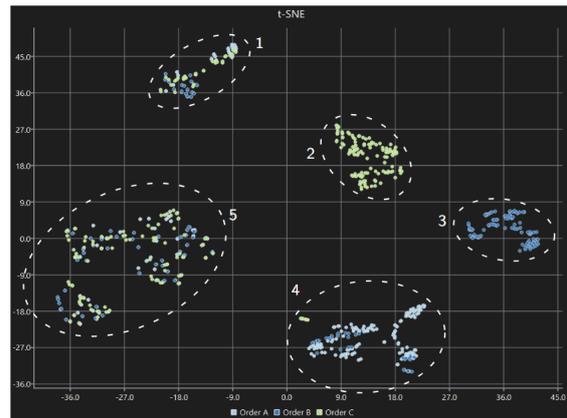


Figure 9: t-SNE for Scenario 2. Five clusters can be retrieved, and are highlighted using dashed ellipses and identified via indices.

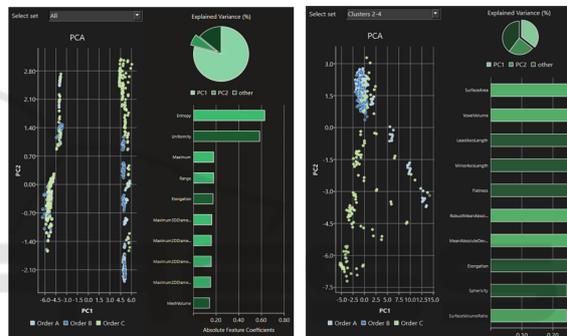


Figure 10: PCAs for Scenario 2: (*left*) PCAs between all orders (A, B, and C), where artifact pipelines without salt noise are on the left-hand side and pipelines with salt noise on the right-hand side; (*right*) PCAs of artifact pipelines containing salt noise from all orders.

**Performance.** The generation of the artifact-affected CT images took 1 hour 57 minutes, most of which was used for data compression and writing. On the other hand, feature extraction took only 37 minutes. The amount of time PCA and t-SNE needed was also negligible for this example scenario since we had a similar number of feature values as in Scenario 1. The relatively large amount of time spent on compression and writing was necessary because the images have an anisotropic resolution of 256 pixels along each axis. Therefore, the file size of one image was around 100.6 MB, which would have resulted in 76 GB of data but were compressed to 35.4 GB.

## 5 DISCUSSION

Flexibility is a strength of our application that is demonstrated in both use cases. Users are empowered to investigate arbitrary artifact configurations,

and they may tailor all settings to the imaging tasks of their interest. As already mentioned in Sec. 2, there is limited prior work similar to ours. Therefore, our results are challenging to contextualize within findings from prior research. However, the fact that the results of our data analysis agree with literature expectations or are supported by reasonable explanations is already a positive indication.

Our approach to evaluating CT artifacts has several limitations. While the framework effectively identifies and quantifies standard artifacts, it may not adequately capture complex artifacts or interactions between them. For example, compound effects with the segmentation step are not currently addressed, as our approach focuses only on the acquisition uncertainties and their propagation through the pipeline. In the future, it would be valuable to increase the selection of segmentation algorithms and explore the impact of this choice (as well as its interplay with the other steps). Additionally, the extension of this methodology to other imaging modalities is theoretically possible. Yet, it presents challenges due to the unique characteristics and types of artifacts associated with each modality. Moreover, the actual impact of the findings discussed in Sec. 4 on clinical decision-making needs a more concrete assessment. Keeping in mind that the investigated uncertainties are not the only ones potentially present in an analytical process (Gillmann et al., 2023; Sacha et al., 2016), a comprehensive study that includes also the other types would be necessary. As a future direction, the framework could integrate a differential analysis feature to compare selected points in the projected views with their respective 3D representations. While the quantitative data and visualizations presented in the use cases are valuable, evaluating the effectiveness of our approach in real-world settings with practitioners would be the next step in understanding the practical implications of our work. For instance, it would be valuable to evaluate the framework in a setting in which the user learns how uncertainties are propagated through the medical imaging pipeline.

## 6 CONCLUSIONS

This work investigated the effects of CT acquisition artifacts on the subsequent steps and outcomes of the medical imaging pipeline. Our results indicate that:

- The uncertainty introduced by CT artifacts impacts segmentation outcomes and feature extraction—varying by artifact type, order, and magnitude.
- Our application supports the effective exploration

of the impact of uncertainties on the outcomes of the medical imaging pipeline.

Our current limitations comprise the choice (and unexplored effect on the pipeline outcome) of segmentation algorithms and the complexity of artifact modeling. Addressing these limitations opens interesting directions for future work that are expected to improve the framework’s generalizability and provide additional insights into uncertainty propagation.

## REFERENCES

- Abdi, H. and Williams, L. J. (2010). Principal component analysis. *WIREs Computational Statistics*, 2(4):433–459.
- Athanasίου, L. S., Rigas, G., Sakellarios, A., Bourantas, C. V., Stefanou, K., Fotiou, E., Exarchos, T. P., Siogkas, P., Naka, K. K., Parodi, O., Vozzi, F., Teng, Z., Young, V. E., Gillard, J. H., Prati, F., Michalis, L. K., and Fotiadis, D. I. (2015). Error propagation in the characterization of atheromatic plaque types based on imaging. *Computer Methods and Programs in Biomedicine*, 121(3):161–174.
- Barrett, J. F. and Keat, N. (2004). Artifacts in ct: Recognition and avoidance. *RadioGraphics*, 24(6):1679–1691. PMID: 15537976.
- Behrens, T., Woolrich, M., Jenkinson, M., Johansen-Berg, H., Nunes, R., Clare, S., Matthews, P., Brady, M., and Smith, S. (2003). Characterization and propagation of uncertainty in diffusion-weighted mr imaging. *Magnetic resonance in medicine : official journal of the Society of Magnetic Resonance in Medicine / Society of Magnetic Resonance in Medicine*, 50:1077–88.
- Boas, F. E. and Fleischmann, D. (2012). Ct artifacts: Causes and reduction techniques. *Imaging in Medicine*, 4:229–240.
- Brodie, K., Osorio, R. S. A., and Lopes, A. (2012). A review of uncertainty in data visualization. In *Expanding the Frontiers of Visual Analytics and Visualization*, pages 81–109.
- Diwakar, M. and Kumar, M. (2018). A review on ct image noise and its denoising. *Biomedical Signal Processing and Control*, 42:73–88.
- Feiner, L. F., Menten, M. J., Hammernik, K., Hager, P., Huang, W., Rueckert, D., Braren, R. F., and Kaissis, G. (2023). *Propagation and Attribution of Uncertainty in Medical Imaging Pipelines*, page 1–11. Springer Nature Switzerland.
- Gillies, R. J., Kinahan, P. E., and Hricak, H. (2016). Radiomics: images are more than pictures, they are data. *Radiology*, 278(2):563–577.
- Gillmann, C., Arbeláez, P., Hernández, J. T., Hagen, H., and Wischgoll, T. (2017). Intuitive Error Space Exploration of Medical Image Data in Clinical Daily Routine. In Kozlikova, B., Schreck, T., and Wischgoll, T., editors, *EuroVis 2017 - Short Papers*. The Eurographics Association.

- Gillmann, C., Maack, R. G. C., Raith, F., Pérez, J. F., and Scheuermann, G. (2023). A taxonomy of uncertainty events in visual analytics. *IEEE Computer Graphics and Applications*, 43(5):62–71.
- Gillmann, C., Saur, D., Wischgoll, T., and Scheuermann, G. (2021). Uncertainty-aware visualization in medical imaging - a survey. *Computer Graphics Forum*, 40(3):665–689.
- Gravel, P., Beaudoin, G., and de Guise, J. A. (2004). A method for modeling noise in medical images. *IEEE Transactions on Medical Imaging*, 23:1221–1232.
- Howard, M., Luttmann, A., and Fowler, M. (2014). Sampling-based uncertainty quantification in deconvolution of x-ray radiographs. *Journal of Computational and Applied Mathematics*, 270:43–51. Fourth International Conference on Finite Element Methods in Engineering and Sciences (FEMTEC 2013).
- Klacansky, P. (2017). Open scivis datasets. <https://klacansky.com/open-scivis-datasets/>.
- Laidlaw, D. H., Trumbore, W. B., and Hughes, J. F. (1986). Constructive solid geometry for polyhedral objects. In *Proceedings of the 13th Annual Conference on Computer Graphics and Interactive Techniques*, SIGGRAPH '86, page 161–170, New York, NY, USA. Association for Computing Machinery.
- Lorensen, W. E. and Cline, H. E. (1987). Marching cubes: A high resolution 3d surface construction algorithm. In *Proceedings of the 14th Annual Conference on Computer Graphics and Interactive Techniques*, SIGGRAPH '87, page 163–169, New York, NY, USA. Association for Computing Machinery.
- Lu, C.-T., Chen, M.-Y., Shen, J.-H., Wang, L.-L., and Hsu, C.-C. (2018). Removal of salt-and-pepper noise for x-ray bio-images using pixel-variation gain factors. *Computers & Electrical Engineering*, 71:862–876.
- Madhura, J. J. and Babu, D. R. R. (2017). An effective hybrid filter for the removal of gaussian-impulsive noise in computed tomography images. In *2017 International Conference on Advances in Computing, Communications and Informatics (ICACCI)*, pages 1815–1820.
- Otsu, N. (1979). A threshold selection method from gray-level histograms. *IEEE Transactions on Systems, Man, and Cybernetics*, 9(1):62–66.
- Potter, K., Rosen, P., and Johnson, C. R. (2012). From quantification to visualization: A taxonomy of uncertainty visualization approaches. In Dienstfrey, A. M. and Boisvert, R. F., editors, *Uncertainty Quantification in Scientific Computing*, pages 226–249, Berlin, Heidelberg. Springer Berlin Heidelberg.
- Preim, B. and Botha, C. P. (2013). *Visual Computing for Medicine: Theory, Algorithms, and Applications*. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 2 edition.
- Raidou, R. G. (2018). Uncertainty visualization: Recent developments and future challenges in prostate cancer radiotherapy planning. In *EuroRV3 2018: EuroVis Workshop on Reproducibility, Verification, and Validation in Visualization*, pages 13–17. The Eurographics Association.
- Reiter, O., Breeuwer, M., Gröller, E., and Raidou, R. (2018). Comparative visual analysis of pelvic organ segmentations. pages 037–041.
- Ristovski, G., Preusser, T., Hahn, H. K., and Linsen, L. (2014). Uncertainty in medical visualization: Towards a taxonomy. *Computers & Graphics*, 39:60–73.
- Roy, C. J. and Oberkampf, W. L. (2011). A comprehensive framework for verification, validation, and uncertainty quantification in scientific computing. *Computer Methods in Applied Mechanics and Engineering*, 200(25):2131–2144.
- Sacha, D., Senaratne, H., Kwon, B. C., Ellis, G., and Keim, D. A. (2016). The role of uncertainty, awareness, and trust in visual analytics. *IEEE Transactions on Visualization and Computer Graphics*, 22(1):240–249.
- Schlachter, M., Raidou, R. G., Muren, L. P., Preim, B., Putora, P. M., and Bühler, K. (2019). State-of-the-art report: Visual computing in radiation therapy planning. In *Computer Graphics Forum*, volume 38, pages 753–779. Wiley Online Library.
- Tian, X. and Samei, E. (2016). Accurate assessment and prediction of noise in clinical ct images. *Medical Physics*, 43(1):475–482.
- van der Maaten, L. (2014). Accelerating t-sne using tree-based algorithms. *Journal of Machine Learning Research*, 15(93):3221–3245.
- van der Maaten, L. and Hinton, G. (2008). Visualizing data using t-sne. *Journal of Machine Learning Research*, 9(86):2579–2605.
- Wu, Y., Yuan, G.-X., and Ma, K.-L. (2012). Visualizing flow of uncertainty through analytical processes. *IEEE Transactions on Visualization and Computer Graphics*, 18:2526–2535.
- Yucel-Finn, A., Mckiddie, F., Prescott, S., and Griffiths, R. (2023). *A Brief History of Time: From the Big Bang to Black Holes*. Elsevier, 1 edition.
- Zhang, J. (2021). Modern monte carlo methods for efficient uncertainty quantification and propagation: A survey. *WIREs Computational Statistics*, 13(5):e1539.
- Zhou, S. K., Greenspan, H., Davatzikos, C., Duncan, J. S., Van Ginneken, B., Madabhushi, A., Prince, J. L., Rueckert, D., and Summers, R. M. (2021). A review of deep learning in medical imaging: Imaging traits, technology trends, case studies with progress highlights, and future promises. *Proceedings of the IEEE*, 109(5):820–838.

## APPENDIX

The symbols  $\mathbf{x} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$  and  $\mathbf{x}_2 = \begin{bmatrix} x \\ y \end{bmatrix}$  denote the image position in all artifact models described below.

### Salt and Pepper Noise Artifact

For the salt and pepper noise artifact, we use the following four parameters:

- $i_s$  radiodensity of salt-pixels
- $i_p$  radiodensity of pepper-pixels
- $p_s$  relative amount of salt-pixels
- $p_p$  relative amount of pepper-pixels

where  $p_s + p_p \leq 1.0$ . The output for all values of the image is then defined as

$$I_{\text{out}}(\mathbf{x}, \Omega) = \begin{cases} i_s & \text{if } \Omega = 1 \\ i_p & \text{if } \Omega = 2 \\ I_{\text{in}}(\mathbf{x}) & \text{if } \Omega = 3 \end{cases} \quad (1)$$

with  $\Omega \sim \text{Categorical}(p_s, p_p, 1.0 - (p_s + p_p))$ .

### Gaussian Noise Artifact

The parameters of Gaussian noise artifacts, following a Poisson distribution, are:

- $\mu$  mean of Gaussian distribution
- $\sigma$  standard deviation of Gaussian distribution

for the output image

$$I_{\text{out}}(\mathbf{x}, T) = I_{\text{in}}(\mathbf{x}) + T \quad (2)$$

with  $T \sim \mathcal{N}(\mu, \sigma^2)$ .

### Cupping Artifact

Our cupping artifacts model uses the following parameters:

- $k_{I_{\text{min}}}$  minimum relative radiodensity at center
- $\mathbf{c} = \begin{bmatrix} c_x \\ c_y \end{bmatrix}$  center voxel with maximum shading

Further, we assume that the shading effect increases linearly towards the center of the artifact to arrive at the model

$$I_{\text{out}}(\mathbf{x}) = I_{\text{in}}(\mathbf{x}) \left( k_{I_{\text{min}}} + (1.0 - k_{I_{\text{min}}}) \frac{d_{\text{max}} - d_c(\mathbf{x}_2)}{d_{\text{max}}} \right) \quad (3)$$

where  $d_c(\mathbf{x}_2) = |\mathbf{x}_2 - \mathbf{c}|$  is the distance from the center, and  $d_{\text{max}} = \max_{\mathbf{x}_2 \in A} d_c(\mathbf{x}_2)$  is the maximum distance from the center to any voxel in the image.

### Ring Artifact

The ring artifacts model takes the following parameters:

- $r$  inner radius of the ring
- $w$  width of the ring
- $k_I$  relative radiodensity of the ring voxels

$$\mathbf{c} = \begin{bmatrix} c_x \\ c_y \end{bmatrix} \quad \text{ring center}$$

The output radiodensities are determined by:

$$I_{\text{out}}(\mathbf{x}) = \begin{cases} k_I I_{\text{in}}(\mathbf{x}) & \text{if } |\mathbf{x}_2 - \mathbf{c}| \in [r, r + w] \\ I_{\text{in}}(\mathbf{x}) & \text{else} \end{cases} \quad (4)$$

### Metal Artifact

Metal artifacts modeling makes two central assumptions: (a) There is only one direction of highest attenuation and one direction of lowest attenuation, which is perpendicular to the former, and (b) the shading effect decreases linearly with the distance when moving away from the boundaries of the structure on a certain slice. It takes the following parameters:

- $k_{I_{\text{min}}}$  minimum relative radiodensity adjacent to the structure in direction of maximum attenuation
- $l$  length of effects beginning at structure border
- $\mathbf{u} = \begin{bmatrix} u_x \\ u_y \end{bmatrix}$  direction of highest attenuation

Before we state our model, let the closest point on an implicit surface  $S(\mathbf{x})$  to a given point on the  $z$ -value of the point be

$$P_S(\mathbf{x}) = \arg \min_{Q \in V_{S,z}} |\mathbf{x} - Q| \quad (5)$$

with  $V_{S,z} = \{\mathbf{x} \in V \mid w = z \wedge S(\mathbf{x}) = 0\}$  being the set of points on the curve defined by the cross section of the  $xy$ -plane with the structure  $S$  at the current  $z$ -value.

Then, we determine the output image using

$$I_{\text{out}}(\mathbf{x}, S_\xi) = \begin{cases} I_{\text{in}}(\mathbf{x}) & \text{if } \nexists P_{S_\xi}(\mathbf{x}) \\ I_{\text{in}}(\mathbf{x}) + R_{\text{max},\xi} a_{\text{dis}} a_{\text{dir}} & \text{else} \end{cases} \quad (6)$$

where  $\xi$  is the implicit structure to which the artifact is applied,  $R_{\text{max},\xi}$  is the maximum radiodensity observed within the structure  $\xi$ ,

$$a_{\text{dis}} = \max \left\{ 0.0, 1.0 - \frac{|\delta|}{l} \right\} \quad (7)$$

is the distance weighting factor with  $\delta = \mathbf{x} - P_{S_\xi}$ , and

$$a_{\text{dir}} = k_{I_{\text{max}}} + \left( 1.0 - 2(1.0 - k_{I_{\text{max}}}) \cos^2 \left( \frac{\delta \cdot \mathbf{u}}{|\delta|} \right) \right) \quad (8)$$

is the direction weighting factor.

### Windmill Artifact

Windmill artifacts are modeled with the following parameters:

- $k_{I_{\text{max}}}$  maximum relative radiodensity adjacent to the structure on a bright streak
- $l$  length of effects beginning at structure border
- $\gamma$  angular width of a single bright streak
- $\phi$  rotation per slice

The output image is defined by:

$$I_{\text{out}}(\mathbf{x}, S_\xi) = \begin{cases} I_{\text{in}}(\mathbf{x}) & \text{if } \nexists P_{S_\xi}(\mathbf{x}) \\ I_{\text{in}}(\mathbf{x}) + R_{\text{max},\xi} k_{I_{\text{max}}} a_{\text{dis}} a_{\text{dir}} & \text{else} \end{cases} \quad (9)$$

where  $\xi$ ,  $R_{\text{max},\xi}$ ,  $a_{\text{dis}}$  as well as  $\delta$  conform to the definitions provided previously for the metal artifact and

$$a_{\text{dir}} = \cos^2 \left( \frac{\pi}{\gamma} \text{atan2}(\delta_y, \delta_x) + \phi(z - z_{\text{min}}) + \frac{3}{4}\pi \right) \quad (10)$$

is the direction weighting factor.

### Motion Artifact

Motion artifacts models take the following parameters:

- $k_{I_{\max}}$  relative radiodensity at the difference region
- $\mathbf{A}$  linear transformation matrix
- $r_g$  Gaussian blur kernel radius
- $\sigma$  Gaussian blur kernel standard deviation

The output image is defined by:

$$I_{\text{out}}(\mathbf{x}, S_{\xi}) = \begin{cases} \Gamma_{r_g, \sigma} * I_t(\mathbf{x}) & \text{if } \mathbf{x} \in V_{S_{\xi}, \mathbf{A}} \\ I_{\text{in}}(\mathbf{x}) & \text{else} \end{cases} \quad (11)$$

where  $\xi$  conforms to the definition provided previously,

$$V_{S_{\xi}, \mathbf{A}}(\mathbf{x}) = \left\{ \mathbf{x} \in V \mid S_{\xi}(\mathbf{x}) > 0 \wedge S_{\xi}(\mathbf{A}^{-1}\mathbf{x}) \leq 0 \right\} \quad (12)$$

is the set of points in the transformation region,

$$I_t(\mathbf{x}) = k_{I_{\max}} I_{\text{in}}(\mathbf{A}^{-1}\mathbf{x}) \quad (13)$$

the intensity after transformation, but before blurring, and  $\Gamma_{r_g, \sigma}$  is the Gaussian kernel used for blurring.

SCITEPRESS  
SCIENCE AND TECHNOLOGY PUBLICATIONS