

NeRF 기반 CT 재구성에서 광선 샘플링과 레이어 정규화의 기여도 분석

Analyzing Ray Sampling and Layer Normalization Contributions in NeRF-based CT Reconstruction

추동혁¹, 안하일², 정윤현^{3*}

Donghyuk Choo¹, Haill An², Younhyun Jung^{3*}

요약

CT는 인체 내부의 3차원 해부학적 구조를 정밀하게 파악하는 데 큰 도움을 주지만, 다수의 X선 이미지를 필요로 하기에 환자가 높은 방사선량에 노출된다는 단점이 있다. 이에 따라 적은 수의 이미지만으로 CT 볼륨을 재구성하는 희소 시점 CT의 필요성이 계속해서 증가하고 있다. NeRF의 등장은 이러한 희소 시점 CT 재구성을 보다 효율적으로 수행할 가능성을 제시하였으며, 최근 연구들은 네트워크 구조의 고도화뿐 아니라 광선 샘플링 전략이나 정규화 기법 등 학습 과정의 세부 설계를 통해 성능 향상을 도모하고 있다. 본 연구는 NeRF 기반 CT 재구성에서 광선 샘플링이 재구성 성능에 미치는 영향을 평가하기 위해 동일한 조건에서 다양한 중요도 기반 샘플링을 비교하였고, 그 결과 전역적인 커버리지를 유지하는 광선 선택이 재구성 품질에 중요함이 확인하였다. 또한 다양한 샘플링 환경에서 레이어 정규화를 적용한 실험을 통해, 내부 표현의 스케일을 안정화하는 정규화가 학습 안정성과 최종 재구성 성능 개선에 유의미하게 기여함을 검증하였다.

핵심어 : 희소 시점 CT 재구성, 신경 방사 필드, 광선 샘플링, 레이어 정규화

Abstract

CT enables precise examination of three-dimensional anatomical structures inside the human body, but it also exposes patients to high radiation because it requires many X-ray projections. This has increased the demand for sparse-view CT, which reconstructs volumes from fewer images. NeRF has opened a promising route for efficient sparse-view CT reconstruction, and recent work seeks improvements not only via network

1 School of Computing, Gachon University, Gyeonggi-do, Republic of Korea [Undergraduate Student]
e-mail: choodh105@gachon.ac.kr

2 School of Computing, Gachon University, Gyeonggi-do, Republic of Korea [Graduate Student]
e-mail: xenotic@gachon.ac.kr

3 School of Computing, Gachon University, Gyeonggi-do, Republic of Korea [Professor]
e-mail: younhyun.jung@gachon.ac.kr (Corresponding author)

* 본 과제(결과물)는 2025년도 교육부 및 경기도의 재원으로 경기RISE센터의 지원을 받아 수행된 지역혁신중심 대학지원체계(RISE)의 결과입니다 (2025-RISE-09-A01). 이 성과는 정부(과학기술정보통신부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임(RS-2025-00554526).

Received(April 23, 2026), Review Result(1st: May 17, 2026), Accepted(June 13, 2026), Published(June 30, 2026)



© 2026 The Authors. Published by NCISS.
This is an open access article licensed under the Creative Commons Attribution-NonCommercial 4.0 International License.
To view a copy of this license, visit <http://creativecommons.org/licenses/by-nc/4.0/>.

design but also through training choices such as ray sampling and normalization. Here, we compared several importance-based ray sampling strategies under identical settings and found that preserving global coverage in ray selection is critical for reconstruction quality. We further show that applying layer normalization across sampling conditions stabilizes internal feature scales, improving training stability and final reconstruction performance.

Keyword : sparse-view CT reconstruction, NeRF, ray sampling, Layer Normalization

1. 서론

X선은 높은 투과성을 바탕으로 대상 피사체 내부 구조를 2차원 이미지로 시각화하는 데 효과적이기 때문에 의료 분야에서 비침습적 진단을 위해 널리 사용된다. 그러나 3차원 해부학적 구조가 단일 2차원 평면에 투영된다는 특성 때문에 서로 다른 구조가 중첩되는 문제가 발생하며, 이에 따라 구조 간 경계를 명확히 구분하기 어렵다는 한계가 존재한다. 이러한 제약을 극복하기 위해 도입된 컴퓨터 단층 촬영(Computed Tomography, CT)은 X선을 다양한 각도에서 획득한 뒤 이를 단층 영상으로 재구성함으로써 3차원 볼륨 정보를 제공하고 병변의 위치와 형태를 더욱 정밀하게 파악할 수 있게 하였다. 그럼에도, CT는 다수의 X선 이미지를 반복적으로 촬영해야 하므로 환자가 비교적 높은 방사선량에 노출된다는 문제가 남아 있다.

이러한 이유로 최근 연구들은 제한된 수의 X선 이미지만으로도 고품질 CT 볼륨을 재구성하는 방법을 모색하고 있다. 특히 합성곱 신경망(Convolutional Neural Network, CNN)을 활용한 접근법은 X선에서 국소적 특징을 효과적으로 추출할 수 있어 X선 이미지들과 CT 볼륨 간의 복잡한 대응 관계를 학습하는 데 유리하다 [1]. 그러나 CNN 기반 방법은 충분한 성능을 확보하기 위해서는 대규모의 X선 이미지와 CT 볼륨의 페어(pair) 데이터가 필요하며, 여러 환자 데이터로 학습된 네트워크는 개인별 해부학적 차이를 충분히 반영하지 못해 재구성된 CT가 학습 데이터 분포에 일반화된 형태로 나타나는 한계가 있다.

한편, 컴퓨터 그래픽 분야에서 제안된 신경 방사 필드(Neural radiance Fields, NeRF) [2]는 X선 이미지를 이용한 CT 볼륨 재구성 문제에 새로운 가능성을 제시하고 있다. NeRF는 다수의 2차원 이미지로부터 3차원 방사장을 다층 퍼셉트론(Multi-Layer Perceptron, MLP)을 통해 암시적으로 표현하는 방식이다. 특히 NeRF의 볼륨 렌더링 과정은 X선 투과로 이미지가 생성되는 물리적 모델과 유사하므로, X선 이미지만을 이용하여 조직 밀도를 암시적으로 학습하고 3차원 CT 볼륨을 재구성할 수 있다. 이 접근법은 대규모 페어 데이터에 대한 의존도가 낮으며, 상대적으로 적은 수의 X선 이미지만으로도 환자 맞춤형 3차원 정보를 복원할 수 있다는 장점이 있다.

최근 NeRF 기반 희소 시점 CT 재구성 연구는 단순히 네트워크 구조를 고도화하는 수준을 넘어, 제한된 관측 정보로 인해 ill-posed문제를 완화하기 위한 다양한 설계 요소들로 확장되고 있다. 구체적으로 X선 이미지에서의 광선 샘플링 전략이나 네트워크 정규화 기법은 재구성 품질과 학습

안정성에 큰 영향을 미치는 핵심 요소로 주목받고 있다.

이에 본 연구는 다양한 광선 샘플링(Ray Sampling) 전략을 동일한 실험 조건에서 정량적으로 비교하고, 레이어 정규화(Layer Normalization, LN)와 같은 네트워크 정규화 기법이 NeRF 기반 CT 재구성의 학습 안정성과 최종 재구성 품질에 어떤 영향을 미치는지 체계적으로 평가하고자 한다.

2. 관련 연구

2.1 신경 방사 필드

NeRF [2]는 다수의 2차원 이미지와 이에 대응되는 카메라 정보를 입력받아, 연속적인 3차원 방사장을 암시적으로 학습하는 네트워크 모델이다. NeRF는 3차원 공간 좌표 $X = (x, y, z)$ 와 시점 방향 $V = (\theta, \phi)$ 을 입력으로 받으면 색상 $C = (r, g, b)$ 및 밀도 σ 를 출력하는 함수를 다층 퍼셉트론으로 근사한다. 이와 같이 학습된 네트워크는 3차원 볼륨에 대한 정보를 암시적으로 내포하며 임의의 시점을 입력하면 그 시점에 대한 이미지를 생성할 수 있다.

NeRF의 작동 원리는 특정 시점의 카메라 광선이 3차원 공간을 통과하며 여러 지점을 샘플링하는 광선 투과 기법(Ray Casting) [3]에 기반한다. 이때 각 광선은 경로 상의 샘플링 지점들에서 예측된 색상과 밀도 값을 광선 방향으로 누적하여 최종 색상을 결정한다. 카메라 광선은 광선의 원점 o 에서 방향 d 로 진행하는 $r(t) = o + td$ 로 정의되며, 여기서 t 는 광선 상의 위치를 나타내는 파라미터이다. 광선이 투과하는 유효 구간은 $[t_1, t_2]$ 로 표현되며, 이 구간에서 샘플링된 값들을 바탕으로 광선의 색상 $C(r)$ 은 다음과 같이 계산된다:

$$C(r) = \int_{t_1}^{t_2} T(t) \cdot \sigma(r(t)) \cdot c(r(t), V) \cdot dt \quad (1)$$

여기서 $\sigma(r(t))$ 및 $c(r(t), V)$ 는 카메라 방향 V 와 함께 카메라 광선을 따라 $r(t)$ 지점의 밀도 및 색상을 나타내고, dt 는 각 통합 단계에서 광선이 이동한 차동 거리를 나타낸다. 또한 전송 함수 $T(t)$ 는 광선이 t 까지 도달할 때 누적 감쇠에 의해 결정되는 투과율을 나타내며, 다음과 같이 정의된다.

$$T(t) = \exp\left(-\int_{t_1}^t \sigma(r(s))ds\right) \quad (2)$$

새로운 시점의 이미지는 해당 이미지의 각 픽셀에서 대응하는 카메라 광선 r 을 추적하여 렌더링을 수행하면 생성된다. 식(1)의 연속 적분은 계산 비용이 매우 크고 신경망 학습 과정에서는 직접적으로 다루기 어려워 원본 NeRF와 후속 연구들에서는 비결정론적 층화 샘플링(stratified

sampling)을 통해 이를 근사화하였다. 이 방식은 광선의 유효 구간 $[t_1, t_2]$ 을 N 개의 균등한 구간으로 분할한 뒤 각 구간에서 샘플을 추출하고, 샘플 지점에서 예측된 값을 누적하여 연속 적분을 이산 합 형태로 근사한다. 이에 따라 식(1)은 다음과 같이 근사될 수 있다.

$$\hat{C}(r) = \sum_{i=1}^N a_i T_i c_i, \text{ where } T_i = \exp\left(-\sum_{j=1}^{i-1} \sigma_j \delta_j\right) \quad (3)$$

δ_i 는 샘플 i 에서 샘플 $i+1$ 까지의 거리를 의미한다. (σ_i, c_i) 는 NeRF의 MLP가 광선 r 을 따라 위치한 샘플 지점 $r(t_i)$ 에서 예측한 밀도와 색상이다. 또한 α_i 는 샘플 i 에서의 알파 합성(opacity) 계수로, 다음과 같이 정의된다.

$$a_i = 1 - \exp(-\sigma_i \delta_i) \quad (4)$$

각 픽셀에 대해 제곱 오차 손실을 최소화하도록 MLP의 파라미터를 최적화하며, 전체 이미지에 대한 손실 함수는 다음과 같이 정의된다.

$$L = \sum_{r \in R} \|\hat{C}(r) - C_{gt}(r)\|_2^2 \quad (5)$$

2.2 감쇠필드

자연 이미지가 물체의 반사광을 기록한다면, X선 이미지는 투과된 X선을 기반으로 내부 구조를 투영하기 때문에 근본적으로 차이가 있다 [4][5]. 이러한 투과 과정은 물질의 흡수에 따른 감쇠로 모델링되며, 감쇠 계수 μ 는 광선이 다양한 조직을 통과할 때 흡수되는 X선의 비율을 나타내고 물질의 밀도에 의해 결정된다. 따라서 X선 기반 NeRF에서는 자연 이미지의 색상 대신 3차원 공간 좌표 $X = (x, y, z)$ 에 감쇠 계수 μ 를 매핑하는 감쇠 필드(attenuation field)를 학습한다. 이때 X선 투영 강도는 비어-램버트 법칙 [6]에 따라 다음과 같이 표현된다.

$$I = I_0 e^{-\sum \mu_i \delta_i} \quad (6)$$

여기서 I 는 검출기에 도달한 감쇠 강도, I_0 는 초기 강도, μ_i 는 샘플 지점 i 의 감쇠 계수, δ_i 는 샘플링 단계를 나타낸다.

3. Method

본 연구에서는 NeRF 기반 희소 시점 CT 재구성을 수행하기 위해 동일한 네트워크 구조와 동일

한 학습 변수를 적용한 실험 환경을 구성하였다.

3.1 광선 샘플링(Ray Sampling)

X선 이미지는 3차원 내부 구조가 광선 경로를 따라 중첩되어 나타나기 때문에, 단일 2차원 이미지 내에 다양한 해부학적 구조와 감쇠 특성이 복합적으로 포함된다. 이러한 특성으로 인해 X선 이미지의 정보는 공간적으로 균일하게 분포하지 않고, 구조적 경계나 급격한 감쇠 변화 등 재구성에 중요한 단서가 특정 영역에 집중될 가능성이 있다. 따라서 우리는 정보량이 높은 구간 또는 중요한 영역에 광선을 상대적으로 더 배치하는 샘플링 전략이 학습 효율과 재구성 품질 향상에 기여할 것이라고 가정한다. 이를 검증하기 위해 다음과 같은 샘플링 전략을 적용하였다.

Baseline으로 둔 Uniform ray sampling은 가장 기본적인 광선 샘플링 방식으로, 학습 단계마다 전체 광선 집합에서 일정 수의 광선을 균일 확률로 무작위 선택하는 전략이다. 이 방식은 투영 영상 내 정보의 공간적 분포를 고려하지 않는 순수 랜덤 샘플링으로 데이터 전반을 편향 없이 탐색할 수 있다는 장점이 있다. 본 연구는 이후에 소개될 중요도 기반 샘플링 전략들과의 공정한 비교를 위해 Uniform ray sampling을 기준 샘플링으로 사용한다.

Gradient 기반 ray sampling은 X선 이미지에서 구조 변화가 큰 경계나 윤곽 영역이 재구성에 더 중요한 단서를 제공할 것이라는 가정에 기반한 전략이다. 이를 위해 각 픽셀에 대해 X선 이미지의 기울기 크기를 계산하고, 그 값을 해당 픽셀에 대응되는 광선의 가중치로 부여한다. 이후 가중치에 비례하도록 광선 선택 확률을 조정함으로써, 기울기가 큰 영역의 광선이 더 자주 샘플링되도록 유도하며, 그 결과 학습 과정에서 해당 영역에 대한 관측과 업데이트가 상대적으로 더 많이 이루어지도록 한다.

MLG ray sampling은 Y. Cai [7]가 제안한 광선 샘플링 방법이다. 이 방식은 우선 X선 이미지에서 의미 있는 영역을 이진 마스크로 분리한 뒤, 해당 영역의 일부는 패치 단위로 선택하여 광선을 집중적으로 샘플링함으로써 지역적 정보를 강화하고, 나머지 영역에서는 광선을 무작위로 샘플링하여 전역적 정보를 함께 학습하도록 한다. 즉, 중요 영역에서 지역적, 전역적 정보를 동시에 반영해 중첩된 복잡한 구조를 보다 효과적으로 학습할 수 있도록 설계된 방법이다.

Local Variance Contrast ray sampling(LVC)은 X선 이미지에서 국소 영역의 구조적 변화가 큰 부분을 우선적으로 학습하기 위한 전략이다. 이 방식은 전체 이미지에 대하여 5x5 커널을 이용해 국소 분산을 계산하고 이를 기반으로 이미지 전역의 분산 가중치 맵을 생성한다. 이를 통해 분산 값의 대비가 클수록 구조 경계나 감쇠 변화가 집중되었을 가능성이 높다고 판단하고 이 영역을 중심으로 더 자주 샘플링되도록 하는 방식이다.

3.2 네트워크 정규화(Network Normalization)

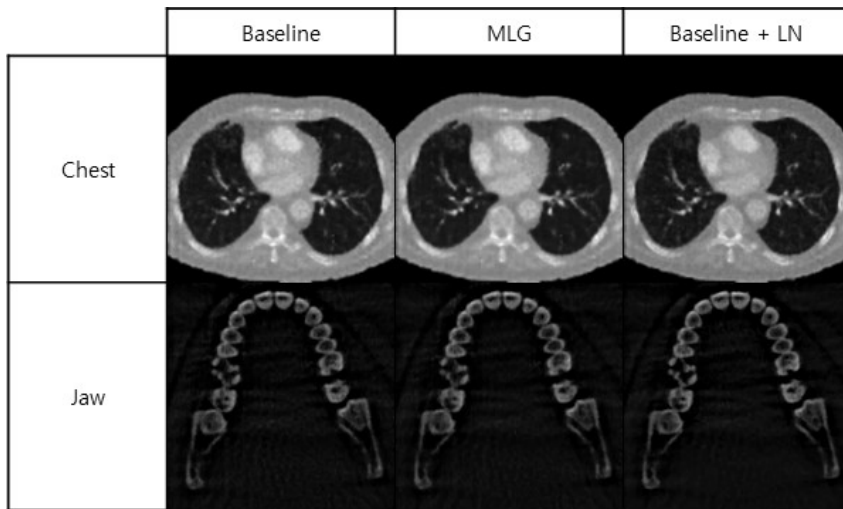
NeRF의 희소 시점 CT 재구성은 관측 정보가 제한되기 때문에, 학습 과정에서 내부 표현의 스케일이 쉽게 불안정해지고, 이는 gradient 분산 증가 및 수렴 불안정으로 이어질 수 있다. 이러한 문제를 완화하기 위해 본 연구는 정규화 적용 여부에 따른 영향을 비교하였다.

레이어 정규화(Layer Normalization, LN)는 각 입력 샘플마다 은닉층의 활성화값 분포를 정규화해 내부 표현의 스케일을 안정화하는 기법이다. 이를 통해 희소 시점 CT 재구성에서 관측 부족으로 인해 표현 스케일이 쉽게 흔들리는 기존 문제를 완화하고, 안정적인 학습을 유도할 수 있다.

4. 실험 설계 및 평가 지표

실험은 PyTorch를 기반으로 설계되었다. 데이터셋으로는 Chest와 Jaw를 사용했으며 입력으로 50개의 X선 이미지를 주었다. 모든 실험은 1500 epoch을 돌렸으며, optimizer는 Adam optimizer를 사용하고 learning rate는 1×10^{-3} 에서 1×10^{-4} 로 점진적으로 감소시켰다. 각 이미지에서 학습에 사용하는 광선의 수는 1024이고, 각 광선은 192개의 포인트 샘플링을 진행한다. 실험의 일관성을 위해 네트워크 아키텍처는 NAF로 고정하였고, 광선 샘플링 전략과 정규화 기법만을 독립 변수로 설정하여 그 효과를 비교하였다.

4.1 광선 샘플링의 영향



[그림 1] Baseline, MLG, LN을 적용한 정성적 평가

[Fig. 1] Qualitative comparison results for the baseline, MLG ray sampling, and LN

[그림 1]은 기존 모델에 광선 샘플링과 네트워크 정규화를 적용한 모델의 실험 결과에 관한 그림이다. 해당 그림을 보면 광선 샘플링을 적용한 경우 특정 부분에서 오히려 학습이 잘 안되는 모

습을 보인다. 또한, Chest와 Jaw 데이터셋에 서로 다른 광선 샘플링 전략을 도입했을 때의 성능 지표를 [표 1]과 같이 정리하였다. 광선 샘플링 실험 결과, MLG 샘플링을 제외한 다른 샘플링 전략들은 baseline인 Uniform 샘플링에 비해 CT 재구성 지표(PSNR, SSIM)가 오히려 낮게 나타났다. 이는 Uniform 및 MLG 샘플링이 광선을 비교적 고르게 선택함으로써 전역적인 정보를 균형 있게 학습할 수 있었던 반면, 나머지 전략들은 특정 광선에 과도하게 집중되면서 전역 정보가 충분히 반영되지 못했기 때문으로 해석된다. 결과적으로 NeRF 기반 CT 재구성에서는 국소적 중요도에 기반한 편향적 샘플링이 항상 성능 향상으로 이어지지 않으며, 전역적 커버리지를 유지하는 샘플링이 더 안정적인 재구성에 기여할 수 있음을 시사한다.

[표 1] 다양한 광선 샘플링을 적용한 PSNR/SSIM 수치

[Table 1] PSNR/SSIM measurements of different ray sampling

Method	Uniform		Gradient		MLG		LVC	
Scene	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
Chest	32.50	0.9596	32.05	0.9537	32.53	0.9595	32.31	0.9558
Jaw	33.99	0.9344	33.29	0.9220	34.02	0.9357	33.68	0.9286

4.2 네트워크 정규화의 영향

다양한 광선 샘플링 환경에서 정규화 방식만을 변경하여 LN의 성능 영향을 비교한 결과를 [표 2]와 같이 정리하였다. [표 1]과 대비했을 때, LN을 적용한 설정에서 PSNR과 SSIM이 전반적으로 향상되는 것을 확인할 수 있다. 이는 관측 부족으로 인해 불안정해지기 쉬운 내부 표현의 스케일을 LN이 효과적으로 제어할 수 있음을 보여주며, 그 결과 학습 안정성과 재구성 품질을 전반적으로 끌어올렸음을 보여준다. 또한, [그림 1]에서 기존 방식보다 정규화를 사용한 이미지에서 노이즈가 적어짐을 확인할 수 있다. 다시 말해, 제한된 투영 정보로 감쇠 필드를 추정해야 하는 NeRF 기반 CT 재구성에서는 학습 표현의 스케일을 안정화하는 것이 재구성 성능과 직결되는 핵심 요소이며, LN은 이러한 안정화를 통해 신뢰도 높은 학습과 일관된 재구성을 뒷받침한다.

[표 2] 다양한 광선 샘플링과 레이어 정규화를 적용한 PSNR/SSIM 수치

[Table 2] PSNR/SSIM measurements of different ray sampling and layer normalization

Method	Uniform+LN		Gradient+LN		MLG+LN		LVC+LN	
Scene	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
Chest	33.40	0.9656	33.27	0.9633	33.99	0.9695	33.77	0.9675
Jaw	35.45	0.9521	35.03	0.9473	35.41	0.9518	35.23	0.9496

5. 결론

본 연구는 NeRF 기반 희소 시점 CT 재구성에서 광선 샘플링 전략과 레이어 정규화의 영향을 동일 조건에서 비교 분석하였다. 실험 결과, 중요 영역에 대한 광선 샘플링 전략이 기존 Uniform sampling보다 성능 향상으로 이어질 것이라 가정했지만, 실제로는 성능 향상이 없거나 감소했으며, 이를 통해 전역적인 커버리지를 유지하는 샘플링의 중요성을 확인했다. 반면, 레이어 정규화는 대부분의 조건에서 학습 안정성과 재구성 품질을 개선해 내부 표현 스케일을 안정화하는 정규화가 CT 재구성에 실질적으로 기여할 수 있음을 보여준다.

References

- [1] P. Henzler, V. Rasche, T. Ropinski, and T. Ritschel, "Single-image tomography: 3D volumes from 2D cranial X-rays," *Computer Graphics Forum*, vol. 37, no. 2, pp. 377-388, May 2018, doi: 10.1111/cgf.13369.
- [2] B. Mildenhall, P. P. Srinivasan, M. Tancik, J. T. Barron, R. Ramamoorthi, and R. Ng, "NeRF: Representing scenes as neural radiance fields for view synthesis," in *16th European Conference on Computer Vision, Virtual*, Aug. 23-28, 2020, pp. 99-106, doi: 10.1145/3503250.
- [3] J. T. Kajiya and B. P. V. Herzen, "Ray tracing volume densities," in *Proceedings of the 11th Annual Conference on Computer Graphics and Interactive Techniques (SIGGRAPH '84)*, Minneapolis, Minnesota, USA, Jul. 23-27, 1984, pp. 165-174, doi: 10.1145/964965.808594.
- [4] A. Molaei et al., "Implicit neural representation in medical imaging: A comparative survey," in *2023 IEEE/CVF International Conference on Computer Vision Workshops (ICCVW)*, Paris, France, Oct. 2-6, 2023, pp. 2373-2383, doi: 10.1109/ICCVW60793.2023.00252.
- [5] R. Zha, Y. Zhang, and H. Li, "NAF: Neural attenuation fields for sparse-view CBCT reconstruction," in *Medical Image Computing and Computer Assisted Intervention - MICCAI 2022*, Singapore, Sep. 18-22, 2022, pp. 442-452, doi: 10.1007/978-3-031-16446-0_42.
- [6] D. F. Swinchart, "The Beer-Lambert law," *Journal of Chemical Education*, vol. 39, no. 7, Art. no. 333, Jul. 1962, doi: 10.1021/ed039p333.
- [7] Y. Cai, J. Wang, A. Yuille, Z. Zhou, and A. Wang, "Structure-aware sparse-view X-ray 3D reconstruction," in *2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Seattle, WA, USA, Jun. 17-21, 2024, pp. 11174-11183, doi: 10.1109/CVPR52733.2024.01062.