

# Self-Calibrated Hyperspectral Neural Radiance Fields for 3D Reconstruction of Bone and Bone Analogues

Neetu Sigger

[N.Sigger@greenwich.ac.uk](mailto:N.Sigger@greenwich.ac.uk)

University of Greenwich

Tuan Thanh Nguyen

University of Greenwich

Sadaf Ashraf

Medway School of Pharmacy

Gianluca Tozzi

University of Greenwich

---

## Article

### Keywords:

**Posted Date:** October 28th, 2026

**DOI:** <https://doi.org/10.21203/rs.3.rs-11346888/v1>

**License:**   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

**Additional Declarations:** No competing interests reported.

---

# Self-Calibrated Hyperspectral Neural Radiance Fields for 3D Reconstruction of Bone and Bone Analogues

Neetu Sigger<sup>1, \*</sup>, Tuan Thanh Nguyen<sup>2</sup>, Sadaf Ashraf<sup>3</sup>, and Gianluca Tozzi<sup>1</sup>

<sup>1</sup>University of Greenwich, School of Engineering, Chatham Maritime, ME44TB, United Kingdom

<sup>2</sup>University of Greenwich, School of Computing & Mathematical Sciences, London, SE109LS, United Kingdom

<sup>3</sup>Medway School of Pharmacy, Chatham Maritime, ME44TB, United Kingdom

\*N.Sigger@greenwich.ac.uk

## ABSTRACT

Hyperspectral imaging (HSI) captures rich, wavelength-dependent information linked to mineralised tissue composition, while 3D reconstruction preserves surface morphology and structural detail. Combining the two offers a powerful, non-contact approach for bone assessment, yet existing hyperspectral 3D reconstruction methods rely on externally estimated camera poses (e.g., via COLMAP), an assumption that is often unreliable on bone surfaces due to their low texture, specular reflectance, and irregular morphology. To address this challenge, we propose BoNeRF-HS, a self-calibrated hyperspectral neural radiance field for 3D reconstruction. BoNeRF-HS jointly optimises camera intrinsics, volume density, and hyperspectral radiance directly from multi-view HSI, removing the need for COLMAP based poses. To exploit the correlation between neighbouring spectral bands, we introduce a gated spectral adapter head that learns a controlled, wavelength-dependent residual correction to the predicted radiance. We evaluate BoNeRF-HS on a newly acquired multi-view hyperspectral dataset (45 views, 120 spectral bands, 400–1000 nm) of mouse bone, trabecular bone analogue, and cortical bone analogue specimens. Experimental results demonstrate that our framework achieves improved reconstruction quality, and better preservation of bone surface details compared with existing approaches. These results establish a practical path toward pose-free, non-contact 3D hyperspectral characterisation of bone, with potential applications in intraoperative visualisation, forensic skeletal analysis, and biomaterial evaluation.

## Introduction

Recent advances in orthopaedic and trauma surgery have increased the use of intraoperative imaging to provide patient specific structural information during bone related procedures. In particular, fluoroscopy and intraoperative three dimensional cone beam computed tomography (CBCT) are widely used for fracture reduction assessment, implant position verification, and surgical navigation<sup>1–3</sup>. These modalities allow surgeons to assess fracture alignment and identify malreduction or implant malposition before wound closure<sup>4,5</sup>. Although these imaging techniques provide valuable structural information, they are associated with several practical limitations, including ionising radiation exposure, two-dimensional projection ambiguity in fluoroscopic imaging, and limited capability to assess tissue specific spectral responses such as wavelength dependent reflectance, oxygenation, or biochemical composition<sup>1,3,6</sup>.

Optical three dimensional reconstruction has therefore been investigated as a non-contact and non-ionising approach for capturing exposed bone surfaces during surgical procedures<sup>7,8</sup>. Structured light imaging<sup>9</sup>, and multi-view reconstruction<sup>8</sup> can provide dense surface geometry and may support surface-based registration, surgical navigation, and intraoperative visualisation<sup>7,8</sup>. However, reliable reconstruction of exposed bone remains challenging because bone surfaces often exhibit limited texture, specular reflections, irregular morphology, and non-uniform illumination, which can affect feature matching, depth estimation, and surface reconstruction accuracy<sup>8</sup>. These same low-texture and specular properties are also the primary reason that feature-based Structure-from-Motion pipelines such as COLMAP are prone to unreliable pose estimation on exposed bone, motivating the self-calibrated formulation adopted in this work. Furthermore, geometric surface models alone do not provide information about tissue composition, spectral response, or surface condition<sup>10,11</sup>.

Hyperspectral imaging (HSI) acquires both spatial and spectral information by capturing a reflectance spectrum at each image pixel<sup>10</sup>. Unlike RGB imaging, which records only three broad colour channels, HSI measures many narrow wavelength bands and can reveal wavelength dependent tissue signatures related to physiology, morphology, and composition<sup>12,13</sup>. This makes HSI particularly attractive for medical and surgical applications, where it may support non-contact and label-free tissue assessment<sup>10,11</sup>. In orthopedic applications, HSI has been explored for several clinically relevant tasks<sup>11,14</sup>. Kistler et al.<sup>14</sup> used HSI to assess knee cartilage and demonstrated that spectral variations could help differentiate damaged cartilage from healthy cartilage. Barberio et al.<sup>11</sup> highlighted the role of intraoperative HSI in supporting tissue discrimination during

surgery, while Chehade et al.<sup>15</sup> evaluated lightweight learning models for HS image segmentation and showed that a multilayer perceptron (MLP) could effectively distinguish bone and cartilage regions, demonstrating its potential in musculoskeletal surgical environments.

Despite these advantages, most surgical HSI approaches analyse the scene as a two-dimensional spectral image<sup>2,10,11</sup>. This limits the interpretation of spectral signatures within the three-dimensional structure of the surgical scene. Combining HSI with 3D reconstruction can overcome this limitation by mapping spectral information onto reconstructed bone surfaces, allowing surface geometry and wavelength-dependent tissue responses to be examined within a common spatial representation<sup>16–18</sup>. Previous studies have shown the feasibility of combining surface reconstruction with hyperspectral or multispectral imaging for surgical applications, including endoscopic surface shape reconstruction<sup>16</sup>, light field HSI<sup>17</sup>, and real-time stereo HSI<sup>18</sup>. Nevertheless, 3D hyperspectral reconstruction of exposed bone surfaces remains relatively underexplored, particularly in settings where accurate camera poses are not externally available.

Neural radiance fields (NeRFs) have emerged as a powerful framework for learning 3D scene representations<sup>19</sup>. A NeRF maps the 3D position and viewing direction to radiance and volume density, and uses differentiable volume rendering to synthesise novel views<sup>19</sup>. Recent work has begun to extend neural rendering from RGB images to hyperspectral data<sup>20,21</sup>. Hyperspectral NeRFs that model wavelength dependent radiance and transmittance<sup>22</sup>, while HS-3D-NeRF uses multi-channel NeRFs for joint 3D surface and hyperspectral reconstruction<sup>21</sup>. Recent hyperspectral Gaussian Splatting<sup>23</sup> methods further explore efficient spectral scene representations for 3D hyperspectral rendering. These methods demonstrate the potential of neural rendering for recovering both geometry and wavelength dependent appearance. However, these hyperspectral 3D methods assume known or externally estimated camera parameters<sup>22,23</sup>. In contrast, self-calibrated RGB NeRF methods such as NeRF<sup>24</sup> jointly optimise camera intrinsics, camera poses, and the radiance field without requiring precomputed poses. Building on this idea, our work extends self-calibrated neural rendering to hyperspectral reconstruction.

This paper presents a novel self-calibrated hyperspectral neural rendering framework, called BoNeRF-HS, for 3D reconstruction of hyperspectral images. BoNeRF-HS jointly optimises camera intrinsics, per-view camera poses, volume density, and hyperspectral radiance, removing the need for COLMAP poses. This work makes three main contributions: first, we construct a multi-view hyperspectral bone dataset containing one mouse bone sample and two bone-analogue samples, which we use to evaluate 3D hyperspectral reconstruction under realistic bone imaging conditions. This dataset is a standalone resource contribution, and is intended to support future benchmarking of hyperspectral 3D reconstruction methods on bone and bone-like structures. Second, we propose BoNeRF-HS, which extends the NeRF<sup>24</sup> self-calibrated formulation from RGB to hyperspectral radiance prediction. The framework jointly optimises camera intrinsics, per-view poses, volume density, and hyperspectral radiance, enabling direct reconstruction of wavelength-dependent appearance from HSI without externally estimated camera poses. Subsequently, to better utilise the spectral information across neighbouring bands, we introduce a gated spectral adapter head within BoNeRF-HS. This module learns a controlled residual correction to the direct spectral prediction, allowing the model to refine spectral responses while preserving the main hyperspectral radiance estimate.

The results show that BoNeRF-HS improves hyperspectral reconstruction quality compared with the other self-calibrated methods, producing sharper rendered views, and better preservation of bone surface details. More broadly, this study demonstrates the potential of self-calibrated hyperspectral neural rendering for non-contact 3D spectral reconstruction of bone surfaces, where geometric structure and spectral responses can be analysed within a unified representation.

## Results

In this section, we present the reconstruction results of the proposed BoNeRF-HS framework and compare them with existing methods using both quantitative metrics and qualitative visualisations.

### Training process

The proposed BoNeRF-HS model was implemented in PyTorch and trained on 120 band HSI. Three specimen types were used in this study, one mouse bone and two bone analogue samples, trabecular bone analogue and the cortical bone analogue. We used approximately 90% of the views for training and the remaining views for testing. The model jointly optimises the hyperspectral radiance field, camera intrinsics, and per-view camera poses. During each training iteration, rays are sampled from the training views, and 128 points are sampled along each ray. Specifically, we use the spectral angle mapper (SAM) loss to preserve spectral-shape similarity between the predicted and ground-truth spectra<sup>25</sup>. The model was trained using the Adam optimiser<sup>26</sup> for 10,000 epochs. During optimisation, randomly sampled  $32 \times 32$  ray batches were used for each training step.

### Performance evaluation

The performance of BoNeRF-HS is evaluated using peak signal-to-noise ratio (PSNR), structural similarity index measure (SSIM), and root mean squared error (RMSE). PSNR and SSIM measure the spatial reconstruction quality of the rendered

**Table 1.** Quantitative comparison across the three hyperspectral samples. Higher PSNR and SSIM indicate better reconstruction quality, while lower RMSE indicate better spectral accuracy. Best results shown in bold.

| Method                  | Mouse bone      |                 |                   | Trabecular bone analogue |                 |                   | Cortical bone analogue |                 |                   |
|-------------------------|-----------------|-----------------|-------------------|--------------------------|-----------------|-------------------|------------------------|-----------------|-------------------|
|                         | PSNR $\uparrow$ | SSIM $\uparrow$ | RMSE $\downarrow$ | PSNR $\uparrow$          | SSIM $\uparrow$ | RMSE $\downarrow$ | PSNR $\uparrow$        | SSIM $\uparrow$ | RMSE $\downarrow$ |
| NeRF <sup>-24</sup>     | 27.87           | 0.8992          | 0.0408            | 23.75                    | 0.7540          | 0.065             | 27.34                  | 0.8992          | 0.0444            |
| SiNeRF <sup>27</sup>    | 22.84           | 0.8439          | 0.0729            | 25.24                    | 0.7864          | <b>0.0547</b>     | 31.23                  | 0.9177          | 0.0274            |
| NoPe-NeRF <sup>28</sup> | 25.83           | 0.8518          | 0.0518            | 24.96                    | 0.7742          | 0.0564            | 30.94                  | 0.9136          | 0.0287            |
| BoNeRF-HS (Ours)        | <b>29.94</b>    | <b>0.9254</b>   | <b>0.0325</b>     | <b>26.13</b>             | <b>0.7919</b>   | <b>0.0547</b>     | <b>31.78</b>           | <b>0.9232</b>   | <b>0.0258</b>     |

views, whereas RMSE assess the spectral accuracy of the reconstructed HSI. All methods are evaluated using the same train test split.

To demonstrate the effectiveness of the proposed BoNeRF-HS framework, we compare it with relevant NeRF based models. NeRF<sup>-24</sup> model is extended from RGB to *B*-band hyperspectral radiance prediction. SiNeRF<sup>27</sup> is included as a sinusoidal neural-field for joint pose and radiance-field optimisation. NoPe-NeRF<sup>28</sup> is also used as a comparison to learn a radiance field without externally provided camera poses. For all comparison methods, the RGB radiance output is replaced by a *B*-band hyperspectral output to enable direct comparison on the same hyperspectral reconstruction task. Table 1 reports the reconstruction results for the mouse bone, trabecular bone analogue, and cortical bone analogue samples. BoNeRF-HS achieves the best PSNR and SSIM on all three samples, indicating improved spatial reconstruction quality compared with NeRF<sup>-24</sup>, SiNeRF<sup>27</sup>, and NoPe-NeRF<sup>28</sup>. On the mouse bone sample, BoNeRF-HS obtains a PSNR of 29.94, an SSIM of 0.9254, and the lowest RMSE of 0.0325. On the trabecular bone analogue, it achieves the highest PSNR and SSIM, while matching the best RMSE of 0.0547. On the cortical bone analogue, BoNeRF-HS again gives the best performance, with a PSNR of 31.78, an SSIM of 0.9232, and the lowest RMSE of 0.0258. These results show that the proposed framework improves hyperspectral reconstruction across both real bone and bone analogue samples.

Figure 1 presents qualitative reconstruction results at spectral band 50 for the mouse bone, trabecular, and cortical bone analogues. BoNeRF-HS is compared with NeRF<sup>-24</sup>, SiNeRF<sup>27</sup>, and NoPe-NeRF<sup>28</sup>. For each sample, the top row shows the rendered HSI band image, while the bottom row shows the absolute difference heatmap with respect to the ground truth. The red boxes mark selected regions of interest (ROI), and the zoomed insets provide enlarged views of these regions to highlight local structural detail, edge sharpness, and texture preservation.

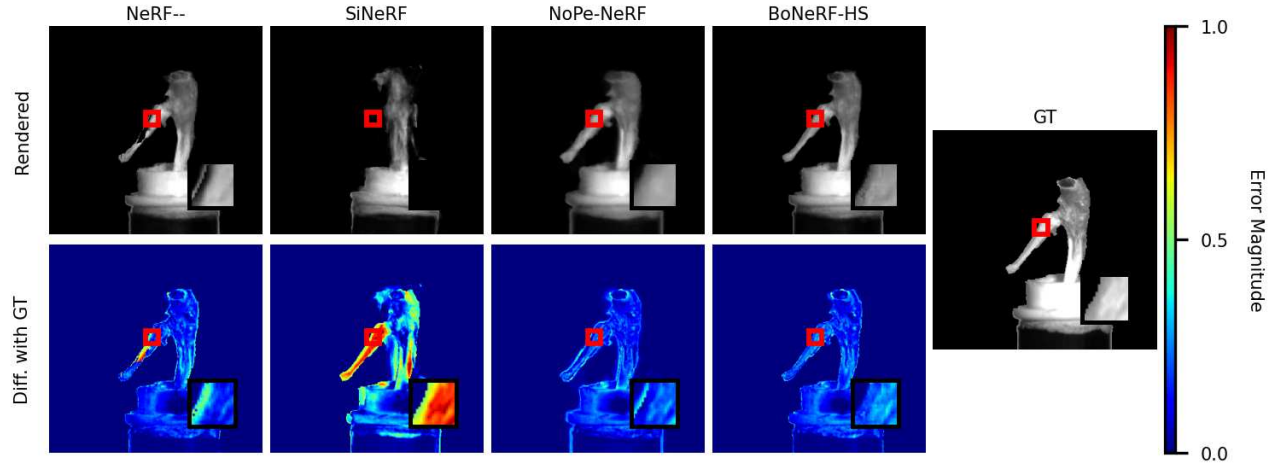
For the mouse sample, the compared methods recover the overall object shape, but their error maps show visible residuals around the bone surface, object boundaries, and fine local structures. SiNeRF<sup>27</sup> produces a smoother appearance but loses some local surface detail, while NeRF<sup>-24</sup> and NoPe-NeRF<sup>28</sup> preserve the global shape but still show noticeable differences from the ground truth. In the zoomed ROI, BoNeRF-HS preserves the selected surface region more clearly and produces lower magnitude errors around the highlighted area.

For the trabecular bone analogue, the porous and textured surface makes reconstruction more challenging. NeRF<sup>-24</sup>, SiNeRF<sup>27</sup>, and NoPe-NeRF<sup>28</sup> reproduce the main cylindrical shape, but the heatmaps show scattered residual errors across the surface texture. The zoomed ROI further shows that local texture and porous surface details are more difficult to preserve. BoNeRF-HS provides a more consistent reconstruction in the highlighted region and reduces the magnitude of the residual errors, indicating improved preservation of local spectral and structural detail.

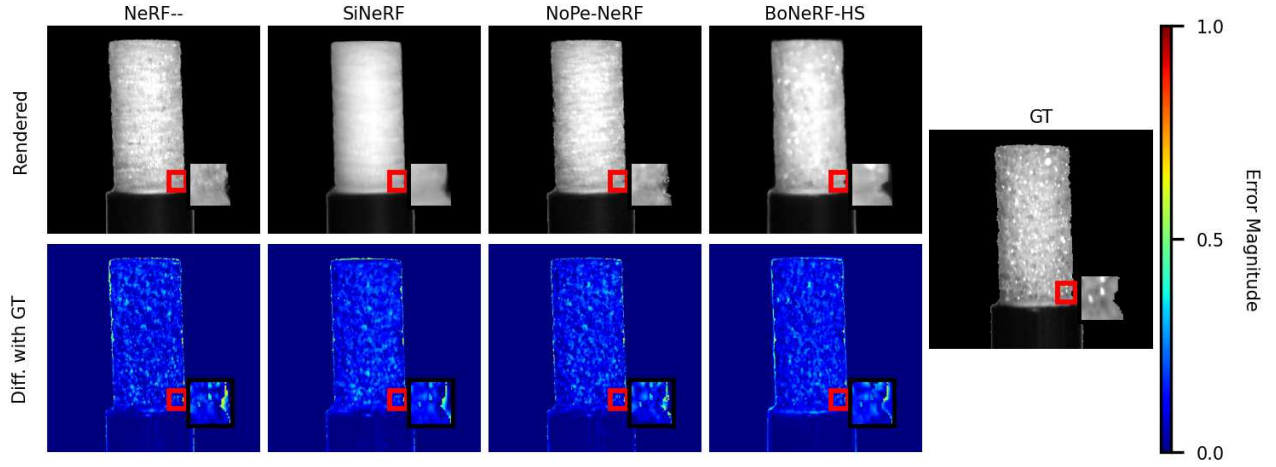
For the cortical bone analogue, all methods reconstruct the compact cylindrical shape. However, the difference heatmaps reveal residual errors along the object boundary and surface edges. NeRF<sup>-24</sup> shows stronger boundary artefacts, while SiNeRF<sup>27</sup> and NoPe-NeRF<sup>28</sup> still retain visible local errors. BoNeRF-HS produces a cleaner reconstruction with lower visible error, in the highlighted ROI and along the compact bone boundary. Overall, the rendered images, heatmaps, and zoomed regions show that BoNeRF-HS better preserves fine surface details and produces more spatially consistent hyperspectral reconstructions across the three samples.

Overall, the rendered images, heatmaps, and zoomed regions show that BoNeRF-HS better preserves fine surface details and produces more spatially consistent hyperspectral reconstructions across the three samples.

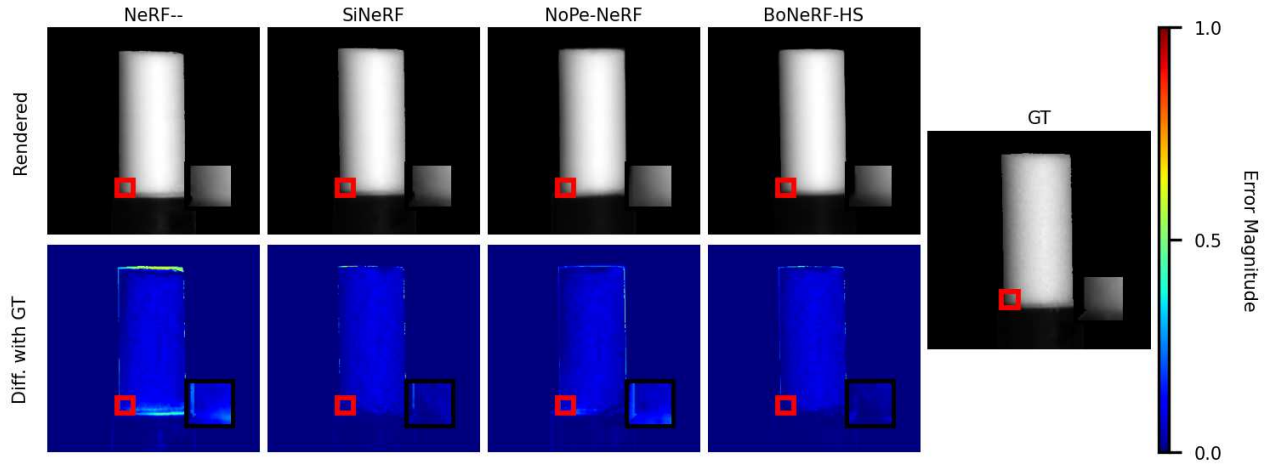
Figure 2 evaluates the spectral reconstruction quality of the compared methods across selected hyperspectral bands (30, 65, and 110) and pixel-wise spectral curves. NeRF<sup>-24</sup>, SiNeRF<sup>27</sup>, and NoPe-NeRF<sup>28</sup> reproduce the general appearance of the samples, but their spectral curves show visible deviations from the ground truth, indicating spectral inconsistency across wavelength bands. BoNeRF-HS produces spectral curves that more closely follow the ground truth, particularly in the near-infrared bands, which are typically more sensitive to noise and lower illumination signal.



(a)



(b)



(c)

**Figure 1.** Qualitative comparison of reconstruction results at spectral band 50 for (a) mouse bone, (b) trabecular bone analogue, and (c) cortical bone analogue. The top row shows reconstructed images from NeRF<sup>24</sup>, SiNeRF<sup>27</sup>, NoPe-NeRF<sup>28</sup>, BoNeRF-HS, and the ground truth. The bottom row shows absolute error maps with respect to the ground truth, where brighter colours correspond to larger errors. The red boxes indicate selected regions of interest, and the zoomed insets provide enlarged views of these regions to highlight local structural details, edge sharpness, and texture preservation.

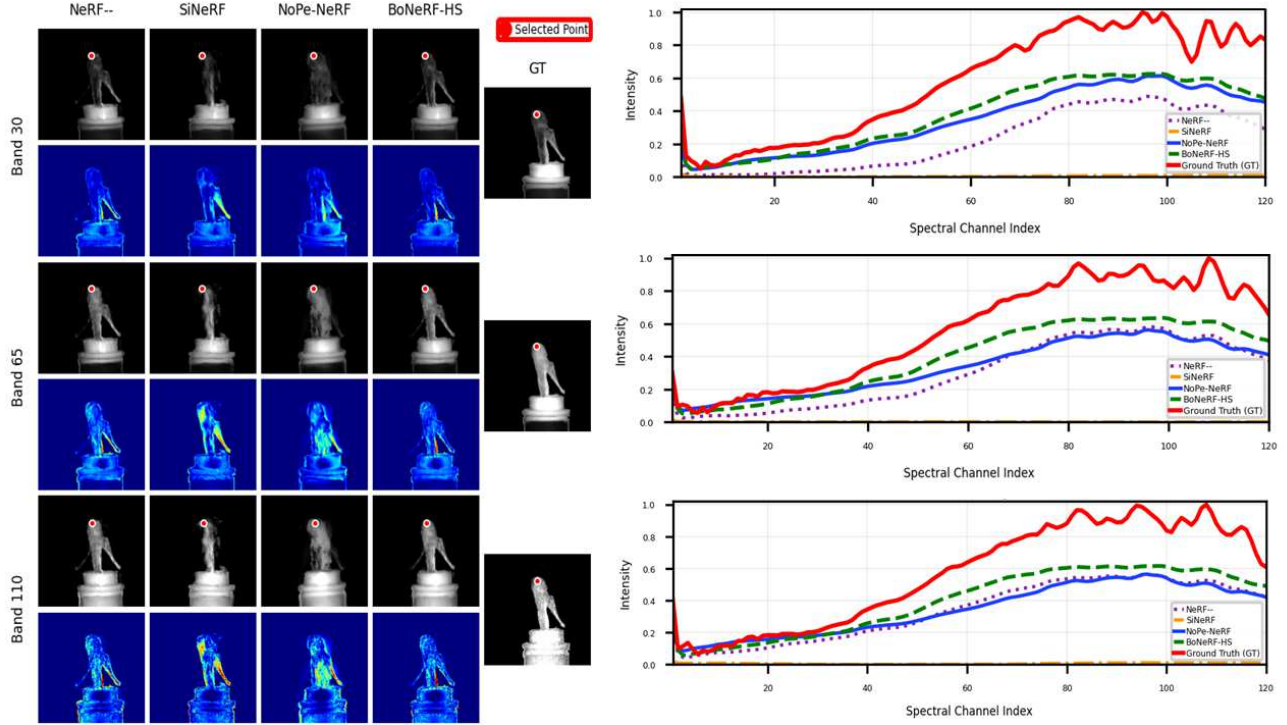


**Figure 2.** Qualitative results on the mouse knee scene across three different bands 30, 65 and 110. The rendered images and difference heatmaps against the ground truth demonstrate the spectral fidelity and spatial consistency of the reconstruction results, particularly under challenging near-infrared and ultraviolet conditions. In addition, we visualize the reconstructed pixel intensities across all the spectral channels of three randomly selected points in the rightmost column. Compared to the baselines, our method exhibits the highest similarity to the ground truth.

## Discussion

In this section, we explore further experiments and discussions on the impact of gated spectral adapter head and the application of the proposed model in soft tissue reconstruction.

### Impact of the Gated Spectral Adapter Head

To analyse the effect of the gated spectral adapter head in the proposed BoNeRF-HS framework, we compare the direct spectral head with three adapter based variants: adapter without gate, adapter with fixed weak gate, and adapter with learned gate as shown in Table 2. In these variants, the adapter acts as a residual spectral correction branch on top of the direct spectral prediction. The adapter without gate applies this correction directly, the fixed weak gate applies it with a constant small strength, and the learned gate adaptively controls how much correction should be added. This ablation evaluates whether residual spectral correction improves the spectral prediction and whether gating is useful for controlling the magnitude of the residual correction. The adapter design ablation shows that residual spectral correction improves the direct spectral head when it is properly controlled: the learned-gate variant improves PSNR by nearly 3 dB over the direct head alone.

The fixed weak gate performs worse than the other adapter variants, suggesting that a single constant correction strength is less flexible than an adaptive gate. We also note that this ablation is averaged across three specimens; per-sample breakdowns and variance across repeated runs would help confirm whether this effect is consistent or driven by a specific sample.

### Performance Analysis on soft tissue

Additional performance analysis was conducted on a mouse tissue sample captured using the same data acquisition and preprocessing procedures described in dataset Acquisition and methods Section. For comparison, BoNeRF-HS was evaluated against NeRF<sup>24</sup>, SiNeRF<sup>27</sup>, and NoPe-NeRF<sup>28</sup>, as reported in Table 3. The proposed BoNeRF-HS model achieved the best overall performance and produced more consistent reconstruction results than the existing approaches. Unlike the bone and bone-analogue results, BoNeRF-HS does not achieve the best SSIM on this sample: NoPe-NeRF<sup>28</sup> scores slightly higher (0.8472 vs. 0.8373). We believe this divergence between PSNR/RMSE and SSIM to the different surface characteristics of soft tissue compared with bone. Soft tissue is more deformable and has lower contrast with locally uniform textures. SSIM which

**Table 2.** Ablation study of the spectral adapter design in BoNeRF-HS, averaged over mouse bone, trabecular and cortical bone analogues. DSH denotes the direct spectral head, SA denotes the spectral adapter, and Gate denotes whether the residual spectral correction is controlled before being added to the direct spectrum. Best results are shown in bold.

| Method                           | DSH | SA | Gate | PSNR $\uparrow$ | SSIM $\uparrow$ | RMSE $\downarrow$ |
|----------------------------------|-----|----|------|-----------------|-----------------|-------------------|
| Direct spectral head             | ✓   | ✗  | ✗    | 26.32           | 0.8428          | 0.0500            |
| Adapter without gate             | ✓   | ✓  | ✗    | 27.37           | 0.8549          | 0.0439            |
| Adapter with fixed weak gate     | ✓   | ✓  | ✓    | 25.90           | 0.8190          | 0.0564            |
| Adapter with learned gate (Ours) | ✓   | ✓  | ✓    | <b>29.29</b>    | <b>0.8810</b>   | <b>0.0376</b>     |

**Table 3.** Quantitative comparison on the mouse tissue sample. Higher PSNR and SSIM indicate better reconstruction quality, while lower RMSE indicates better intensity accuracy.

| Method                  | Mouse bone      |                 |                   |
|-------------------------|-----------------|-----------------|-------------------|
|                         | PSNR $\uparrow$ | SSIM $\uparrow$ | RMSE $\downarrow$ |
| NeRF <sup>24</sup>      | 22.16           | 0.7936          | 0.0783            |
| SiNeRF <sup>27</sup>    | 23.90           | 0.8204          | 0.0647            |
| NoPe-NeRF <sup>28</sup> | 26.47           | <b>0.8472</b>   | 0.0477            |
| <b>BoNeRF-HS (Ours)</b> | <b>27.56</b>    | 0.8373          | <b>0.0425</b>     |

focuses on local structural and contrast patterns, might therefore weigh these features differently than it does pixel-wise intensity errors. This suggests that BoNeRF-HS’s spectral adapter, which is tuned to exploit inter-band correlation, may deliver strong per-pixel and spectral accuracy without necessarily optimising for the local structural contrast that SSIM captures on non-rigid, low-texture soft tissue. This distinction between rigid mineralised bone and deformable soft tissue is worth investigating further, and we flag it as an open question rather than a fully resolved result.

Figure 3 further shows the rendered HSI bands and absolute difference heatmaps for the mouse tissue sample. The compared methods recover the general appearance of the tissue, but their error maps show visible residuals near the boundaries of the tissue, textured regions, and fine local structures. In contrast, BoNeRF-HS produces a sharper rendering that is visually closer to the ground truth and shows lower magnitude errors.

In summary, HSI captures rich wavelength dependent information, which is important for analysing mineralised tissues. In this work, we propose BoNeRF-HS, a self-calibrated hyperspectral neural radiance field for 3D bone reconstruction. In conclusion, most current HSI 3D reconstruction methods rely on externally estimated camera poses or RGB based radiance modelling, which may limit their applicability to multi view hyperspectral acquisition. In contrast, BoNeRF-HS jointly optimises camera intrinsics, volume density, and hyperspectral radiance, allowing accurate reconstruction without COLMAP poses. To better utilise the spectral information present in HSI data, we incorporate a gated spectral adapter head that learns wavelength dependent residual information while preserving the direct spectral prediction. We demonstrated the effectiveness of BoNeRF-HS using multi-view HSI data from mouse bone, trabecular, and cortical bone analogues. The results show that the proposed framework improves hyperspectral reconstruction quality compared with the existing methods, producing sharper rendered views and better preservation of fine bone surface details.

Beyond the reconstruction experiments, BoNeRF-HS may have broader relevance in clinical, forensic, and biomechanical applications. Clinically, the framework could support non-contact 3D spectral visualisation of exposed bone surfaces, where both surface structure and tissue specific spectral response are important for future image guided orthopaedic applications<sup>10,11,16,18</sup>. Since our experiments include mouse tissue, mouse bone, trabecular, and cortical bone analogues, the results provide an initial step towards evaluating self-calibrated 3D hyperspectral reconstruction across both biological tissue and controlled bone like materials.

The framework may also be useful in forensic science, where spectral and shape based analysis of skeletal material has been explored for post-mortem interval estimation<sup>29</sup>, and human remain classification<sup>30,31</sup>. In this context, BoNeRF-HS could provide a non-contact approach for documenting skeletal samples while preserving both surface morphology and wavelength dependent spectral information. Further validation on larger and more diverse tissue, bone, and forensic datasets is required before translation to these applications.

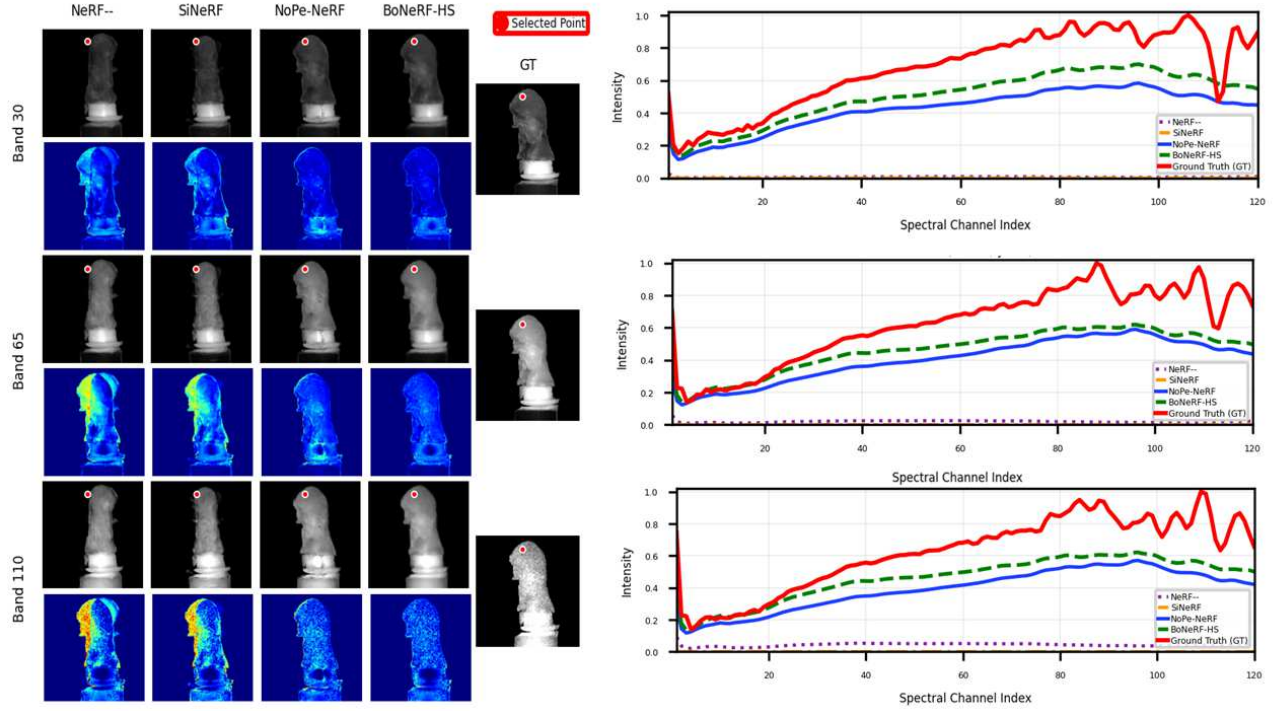


**Figure 3.** Qualitative results on the mouse tissue scene across three different bands 30, 65 and 110. Each column shows renderings and difference maps for NeRF<sup>24</sup>, SiNeRF<sup>27</sup>, NoPe-NeRF<sup>28</sup> and proposed method. In addition, we visualize the reconstructed pixel intensities across all the spectral channels of three randomly selected points in the rightmost column. Compared to the baselines, our method exhibits the highest similarity to the ground truth.

## Limitations and Future Work

The experiments were conducted on a single sample per type (one mouse bone, one trabecular analog, one cortical analog). Therefore, the reported improvements should be interpreted as promising initial evidence rather than a statistically validated benefit across biological variability. The dataset was acquired with a fixed camera, fixed illumination, and a motorised turntable, a benchtop setting that differs from the freehand, dynamic, and variably illuminated conditions of an actual intraoperative environment that motivates this work. Translation to real surgical use will require validation under less controlled acquisition conditions. Bone surfaces are difficult for COLMAP due to low texture and specular reflection, which is the central motivation for self-calibration. This argument is not empirically demonstrated on our dataset (e.g., by showing COLMAP failing or producing high pose error on the acquired views). Future work should include this direct comparison to substantiate the motivation empirically rather than by citation alone. As discussed in Section Discussion, BoNeRF-HS does not achieve the best SSIM on the soft tissue sample despite leading on PSNR and RMSE, indicating that further investigation is needed into how the method performs on deformable, low-texture tissue versus rigid mineralised bone.

## Dataset Acquisition and Methods

In this section, we introduce the hyperspectral dataset used in this study. We then present the proposed BoNeRF-HS framework, a self-calibrated hyperspectral neural radiance field designed for 3D reconstruction.

### Dataset and Preprocessing

We acquire multi-view hyperspectral images using a hyperspectral camera and a turntable motorized rotational sample stage. All hyperspectral data were acquired using a NIREOS HERA VNIR hyperspectral camera<sup>32</sup>. The camera operates in the visible and near-infrared (NIR) of 400–1000nm. In this work, each HSI was exported with 120 spectral bands and spatial resolution of  $320 \times 332$  pixels.

The dataset consists of three specimen types, one mouse bone and two bone analogues, namely a trabecular and a cortical bone analogues. Mouse knee were dissected following euthanasia via an appropriate Schedule 1 technique (as listed by the UK Home Office). The tibiofemoral knee joints were then fixed for 48 hours in 4% paraformaldehyde (PFA) and

washed in phosphate buffered saline, prior to being trimmed to remove all the exterior muscle and expose the underlying bone for imaging. The bone analogues were prepared from Sawbones polyurethane foams (Pacific Research Laboratories, WA, USA), using open-cell PCF15 foam to represent trabecular-like porous bone and closed cell PCF40 foam to represent denser cortical-like bone. Cylindrical samples of 8 mm diameter and 10 mm length were milled from the foam blocks. Trabecular bone is characterised by a porous cancellous architecture formed by interconnected struts and rods, whereas cortical bone is a denser compact structure with much lower porosity<sup>33,34</sup>. The use of bone analogue materials is motivated by previous phantom and biomaterial studies, where synthetic or cast materials are designed to approximate selected physical or imaging properties of bone<sup>35,36</sup>.

For data acquisition, mouse bone and bone analogue specimens were mounted on a sample holder and placed on the motorized rotational sample stage. The camera is kept fixed under constant illumination, while the object is rotated to obtain HS images from multiple viewing directions. Each specimen contains 45 hyperspectral views acquired from 0° to 360° with approximately 8° angular spacing.

The acquired HSI is preprocessed. For reflectance calibration, the raw hyperspectral image  $I_{\text{raw}}$  was corrected using a dark reference  $I_{\text{dark}}$  and a white reference  $I_{\text{white}}$ . The dark reference was captured with the camera lens closed and the illumination switched off, while the white reference was captured using a reference white target. The calibrated reflectance image  $I_{\text{ref}}$  was computed as

$$I_{\text{ref}} = \frac{I_{\text{raw}} - I_{\text{dark}}}{I_{\text{white}} - I_{\text{dark}}} \quad (1)$$

Background removal was applied during preprocessing. Since the camera and illumination were fixed and only the specimen rotated, the background remained stationary with respect to the camera and did not rotate with the object. This could introduce inconsistent background information across views. Therefore, a foreground mask was generated for each hyperspectral view and applied uniformly to all spectral bands. Pixels outside the foreground region were set to zero, producing a pure black background in every wavelength channel. The final processed dataset for each specimen contains 45 calibrated HS views.

## Methods

Figure 4 illustrates the proposed BoNeRF-HS framework. The input to the proposed framework is a set of multi-view HS image captured from different viewpoints of the same specimen. Let  $\mathbf{I}_i \in \mathbb{R}^{H \times W \times B}$  denote the  $i$ -th hyperspectral image cube in the acquired multi-view dataset, where  $H$  and  $W$  denote the image height and width, respectively, and  $B$  denotes the number of spectral bands. The model is trained using the HSI views directly, rather than converting them to RGB images. This allows the network to model the wavelength dependent appearance across all  $B$  spectral bands. Each component of proposed framework describe in this section.

### Self-Calibrated Camera Parameter

Conventional NeRF method assumes known camera intrinsics and extrinsics<sup>19,37</sup>. In contrast, BoNeRF-HS framework does not require externally estimated camera parameters. Inspired by NeRF—<sup>24</sup>, the proposed BoNeRF-HS framework avoids the need for externally provided camera calibration. The camera intrinsics and per view extrinsics are treated as learnable parameters and are optimised together with the hyperspectral radiance field.

For each view  $i$ , the camera intrinsics are parameterised by learnable focal lengths  $f_x$  and  $f_y$ . The intrinsic matrix is defined as

$$K = \begin{bmatrix} f_x & 0 & W/2 \\ 0 & f_y & H/2 \\ 0 & 0 & 1 \end{bmatrix}. \quad (2)$$

The extrinsic parameters of view  $i$  are represented by a camera-to-world transformation

$$T_i = \begin{bmatrix} R_i & t_i \\ 0^\top & 1 \end{bmatrix} \quad (3)$$

where  $R_i$  is the camera rotation and  $t_i$  is the camera translation vector.

Both the camera intrinsics  $(f_x, f_y)$  and the extrinsics  $(R_i, t_i)$  are optimised during training. Therefore, the model does not require externally estimated camera calibration or camera poses.

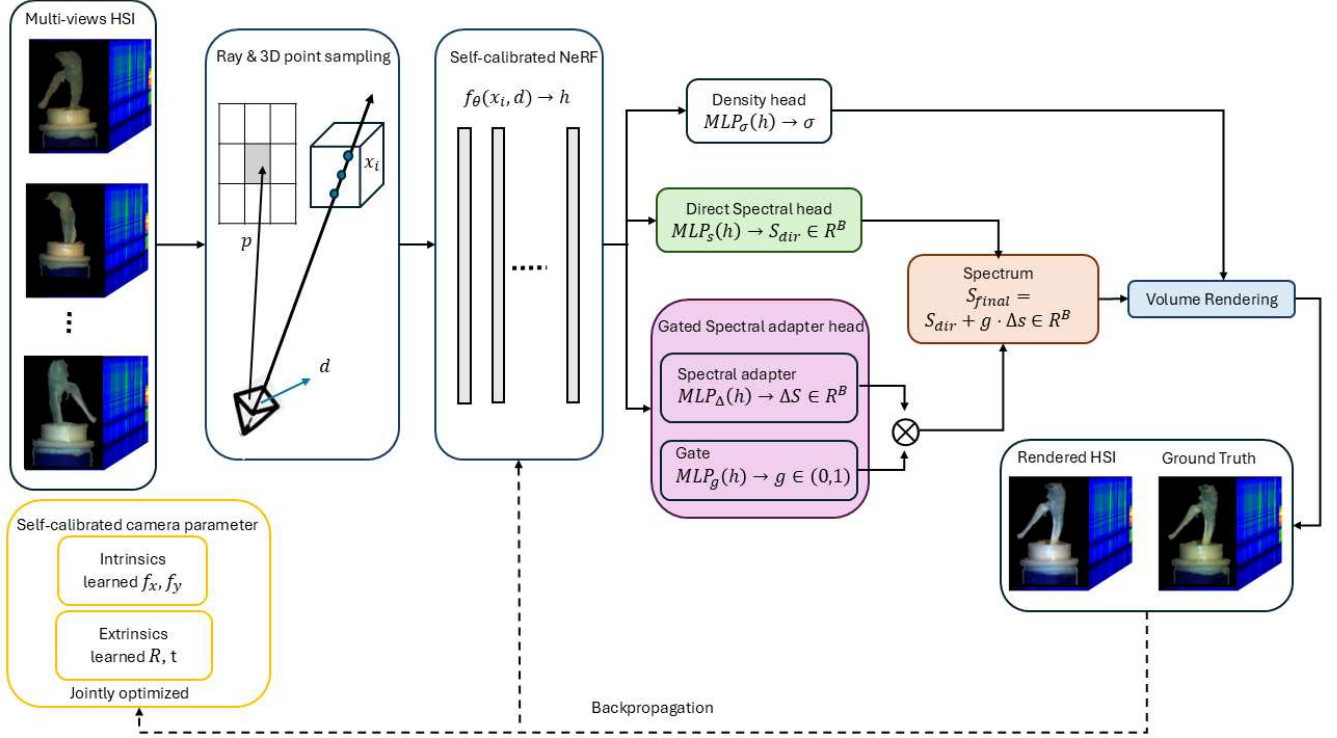


**Figure 4.** Proposed BoNeRF-HS pipeline for self-calibrated hyperspectral 3D reconstruction. Multi-view HS images with spectral bands  $B$  are used as input. Camera intrinsics and extrinsics are jointly optimized with the NeRF model, avoiding the need for externally provided camera poses. The NeRF backbone (vertical bar denote fully connected MLP layers) maps sampled 3D points and viewing directions to a latent feature  $h$ , from which a density head predicts  $\sigma$ , a direct spectral head predicts  $S_{dir}$ , and a gated spectral adapter predicts  $\Delta S$  and  $g$ . The spectrum is  $S_{final} = S_{dir} + g \cdot \Delta S$ , which is volume-rendered to reconstruct the HS images.

### Ray Generation and 3D Point Sampling

For a pixel location  $p$  in image  $i$ , a camera ray  $r$  is generated using the current estimates of the camera intrinsics and extrinsics. The ray is defined as

$$r_{i,p}(t) = o_i + t d_{i,p}, \quad (4)$$

where  $o_i$  is the camera centre,  $d_{i,p}$  is the viewing direction passing through pixel  $p$ , and  $t$  denotes the distance along the ray. A set of  $N_s$  sample points is drawn along each ray

$$x_{i,p,k} = o_i + t_k d_{i,p}, \quad k = 1, \dots, N_s. \quad (5)$$

Each sampled point  $x_{i,p,k}$  and the viewing direction  $d_{i,p}$  are used as inputs to the hyperspectral NeRF model.

### Self-Calibrated Hyperspectral NeRF

The NeRF model is implemented as a multilayer perceptron (MLP)<sup>38</sup>, following the standard NeRF model<sup>19</sup>. Unlike conventional NeRF models that predict RGB colour values, our hyperspectral NeRF is designed to learn a continuous 3D representation that predicts wavelength-dependent spectral responses across all  $B$  spectral bands. The MLP acts as a neural representation of the scene by learning the relationship between 3D spatial locations, viewing directions, and hyperspectral appearance. The sampled 3D point and viewing direction are first encoded using positional encoding,

$$h = f_\theta(\gamma(x_{i,p,k}), \gamma(d_{i,p})), \quad (6)$$

where  $\gamma(\cdot)$  denotes positional encoding,  $f_\theta$  is the NeRF backbone, and  $h$  is the latent feature vector. This shared feature is passed to three prediction heads: a density head, a direct spectral head, and the proposed gated spectral adapter head. Each head has a specific role: the density head models the spatial structure of the scene, the direct spectral head provides the initial hyperspectral prediction, and the gated spectral adapter refines this prediction by learning a spectral residual correction.

## Density Head

The density head predicts the volume density at each sampled 3D point,

$$\sigma = \text{MLP}_\sigma(h). \quad (7)$$

The density  $\sigma$  controls how much each sampled point contributes to the final rendered ray. This follows the volumetric density modelling used in NeRF<sup>19</sup>. Specifically, the density represents the likelihood that a sampled 3D point belongs to the object surface. High-density regions contribute more strongly during rendering, allowing the network to implicitly learn the underlying 3D geometry without requiring an explicit mesh representation.

## Direct Spectral Head

The original NeRF<sup>19</sup> predicts RGB radiance. To support HSI, we replace the RGB colour head with a  $B$ -band spectral head. This is consistent with hyperspectral NeRF model that extend radiance prediction to multi-channel or wavelength dependent spectral outputs<sup>20,39</sup>. The direct spectral head maps the latent feature  $h$  to an initial  $B$  band spectrum,

$$S_{\text{dir}} = \text{MLP}_s(fh), \quad S_{\text{dir}} \in \mathbb{R}^B. \quad (8)$$

Here,  $S_{\text{dir}}$  denotes the direct hyperspectral radiance predicted by the model before the adapter correction. The purpose of this head is to learn the coarse wavelength-dependent appearance of the scene.

## Gated Spectral Adapter Head

Although the direct spectral head predicts a full  $B$  band hyperspectral spectrum, it treats the spectral bands primarily as output channels. HSI contain many contiguous wavelength bands, and neighbouring bands often exhibit strong spectral correlation rather than being independent. Therefore, modelling the spectral output as a structured wavelength dependent signal is important for preserving material specific spectral signatures<sup>40,41</sup>. To better capture these inter-band dependencies, we introduce a gated spectral adapter head after the NeRF backbone. It is inspired by the general idea of lightweight adapter and low-rank model adaptation<sup>42</sup>.

The spectral adapter predicts a residual correction  $\Delta S \in \mathbb{R}^B$ ,

$$\Delta S = \text{MLP}_\Delta(h). \quad (9)$$

In parallel, a gate branch predicts a confidence value  $g \in (0, 1)$ ,

$$g = \text{sigmoid}(\text{MLP}_g(h)). \quad (10)$$

The final spectrum used for rendering is computed as

$$S_{\text{final}} = S_{\text{dir}} + g \cdot \Delta S, \quad S_{\text{final}} \in \mathbb{R}^B. \quad (11)$$

This Eq. 11 allows the adapter to learn a gated wavelength correlated correction. The adapter therefore acts as a refinement branch rather than replacing the direct spectral prediction. By learning only the residual correction, the model can preserve the global spectral structure while improving local wavelength-dependent errors.

## Hyperspectral Volume Rendering

For each sampled point along a ray, the model predicts a density value  $\sigma_k$  and a final hyperspectral spectrum  $S_{\text{final},k}$ . We render the final hyperspectral pixel using the differentiable volume rendering formulation adopted from NeRF<sup>19</sup>. The opacity of sample  $k$  is

$$\alpha_k = 1 - \exp(-\sigma_k \delta_k), \quad (12)$$

where  $\delta_k$  is the distance between adjacent sampled points. The accumulated transmittance is

$$T_k = \prod_{j < k} (1 - \alpha_j), \quad (13)$$

and the rendering weight is

$$w_k = T_k \alpha_k. \quad (14)$$

The rendered hyperspectral pixel is obtained by integrating the  $B$ -band spectra along the ray,

$$\hat{I}_{i,p} = \sum_{k=1}^{N_s} w_k S_{\text{final},k}, \quad \hat{I}_{i,p} \in \mathbb{R}^B. \quad (15)$$

Applying this operation to all pixels produces the rendered hyperspectral image

$$\hat{I}_i \in \mathbb{R}^{H \times W \times B}. \quad (16)$$

## Data availability

The authors declare that all data supporting the current study are available under reasonable request.

## References

1. Richter, P. H., Gebhard, F., Eickhoff, A. & Schütze, K. New advances in intra-operative imaging in trauma. *EFORT Open Rev.* **3**, 168–172, DOI: [10.1302/2058-5241.3.170055](https://doi.org/10.1302/2058-5241.3.170055) (2018).
2. Tonetti, J., Boudissa, M., Kerschbaumer, G. & Seurat, O. Role of 3d intraoperative imaging in orthopedic and trauma surgery. *Orthop. & Traumatol. Surg. & Res.* **106**, S19–S25, DOI: [10.1016/j.otsr.2019.05.021](https://doi.org/10.1016/j.otsr.2019.05.021) (2020).
3. Keil, H. *et al.* Intraoperative revision rates due to three-dimensional imaging in orthopedic trauma surgery: results of a case series of 4721 patients. *Eur. J. Trauma Emerg. Surg.* **49**, 373 – 381 (2022).
4. Luxenhofer, M. *et al.* Diagnostic accuracy of intraoperative ct-imaging in complex articular fractures – a cadaveric study. *Sci. Reports* **10**, 4530, DOI: [10.1038/s41598-020-61267-w](https://doi.org/10.1038/s41598-020-61267-w) (2020).
5. Greeven, A., Van Groningen, J., Schep, N., Van Lieshout, E. & Verhofstad, M. Open reduction and internal fixation versus closed reduction and percutaneous fixation in the treatment of bennett fractures: A systematic review. *Injury* **50**, 1470–1477, DOI: <https://doi.org/10.1016/j.injury.2019.06.027> (2019).
6. Keil, H. & Trapp, O. Fluoroscopic imaging: New advances. *Injury* **53**, S8–S15, DOI: [10.1016/j.injury.2022.05.035](https://doi.org/10.1016/j.injury.2022.05.035) (2022).
7. He, G., Mustahsan, V. M., Bielski, M. R., Kao, I. & Khan, F. A. Report on a novel bone registration method: A rapid, accurate, and radiation-free technique for computer- and robotic-assisted orthopedic surgeries. *J. Orthop.* **23**, 227–232, DOI: [10.1016/j.jor.2021.01.010](https://doi.org/10.1016/j.jor.2021.01.010) (2021).
8. Schierbaum, A., Neiss-Theuerkauff, T., Luhmann, T., Wallhoff, F. & Sieberth, T. Investigations on 3d reconstruction of bones in surgery using a handheld trinocular camera system. *The Int. Arch. Photogramm. Remote. Sens. Spatial Inf. Sci.* **XLVIII-2/W7-2024**, 145–151, DOI: [10.5194/isprs-archives-XLVIII-2-W7-2024-145-2024](https://doi.org/10.5194/isprs-archives-XLVIII-2-W7-2024-145-2024) (2024).
9. Bell, T., Li, B. & Zhang, S. *Structured Light Techniques and Applications* (2016).
10. Lu, G. & Fei, B. Medical hyperspectral imaging: A review. *J. Biomed. Opt.* **19**, 010901, DOI: [10.1117/1.JBO.19.1.010901](https://doi.org/10.1117/1.JBO.19.1.010901) (2014).
11. Barberio, M. *et al.* Intraoperative guidance using hyperspectral imaging: A review for surgeons. *Diagnostics* **11**, 2066, DOI: [10.3390/diagnostics11112066](https://doi.org/10.3390/diagnostics11112066) (2021).
12. Sigger, N., Nguyen, T. & Tozzi, G. Detecting temporal changes in skin surrogate model using hyperspectral imaging. *MRS Commun.* **15**, DOI: [10.1557/s43579-025-00777-y](https://doi.org/10.1557/s43579-025-00777-y) (2025).
13. Barberio, M. *et al.* Intraoperative guidance using hyperspectral imaging: A review for surgeons. *Diagnostics* **11**, DOI: [10.3390/diagnostics11112066](https://doi.org/10.3390/diagnostics11112066) (2021).
14. Kistler, M. *et al.* Intraoperative hyperspectral imaging (hsi) as a new diagnostic tool for the detection of cartilage degeneration. *Sci. Reports* **12**, 608, DOI: [10.1038/s41598-021-04642-5](https://doi.org/10.1038/s41598-021-04642-5) (2022).
15. Hage Chehade, A. *et al.* Hyperspectral imaging for orthopedic surgery: lightweight ai enables real-time bone and cartilage segmentation. *Med. & Biol. Eng. & Comput.* **64**, 1913–1929, DOI: [10.1007/s11517-026-03551-x](https://doi.org/10.1007/s11517-026-03551-x) (2026).
16. Lin, J. *et al.* Dual-modality endoscopic probe for tissue surface shape reconstruction and hyperspectral imaging enabled by deep neural networks. *Med. Image Analysis* **48**, 162–176, DOI: [10.1016/j.media.2018.06.004](https://doi.org/10.1016/j.media.2018.06.004) (2018).
17. Wisotzky, E. L., Hilsmann, A. & Eisert, P. 3d hyperspectral light-field imaging: a first intraoperative implementation. *Curr. Dir. Biomed. Eng.* **9**, 611–614, DOI: [10.1515/cdbme-2023-1153](https://doi.org/10.1515/cdbme-2023-1153) (2023).
18. Wisotzky, E. L. *et al.* Real-time fusion of stereo vision and hyperspectral imaging for objective decision support during surgery. *Comput. Vis. Image Underst.* **262**, 104541, DOI: [10.1016/j.cviu.2025.104541](https://doi.org/10.1016/j.cviu.2025.104541) (2025).
19. Mildenhall, B. *et al.* Nerf: Representing scenes as neural radiance fields for view synthesis. In *European Conference on Computer Vision* (2020).
20. Chen, G. *et al.* Hyperspectral neural radiance fields. *arXiv preprint arXiv:2403.14839* (2024).
21. Ku, K., Jubery, T. Z., Krishnamurthy, A. & Ganapathysubramanian, B. Hs-3d-nerf: 3d surface and hyperspectral reconstruction from stationary hyperspectral images using multi-channel nerfs (2026). [2602.16950](https://arxiv.org/abs/2602.16950).
22. Chen, G. *et al.* Hyperspectral neural radiance fields (2024). [2403.14839](https://arxiv.org/abs/2403.14839).

23. Narayanan, S. K., Zhao, L., Gan, L. & Chen, Y. Diffusion denoised hyperspectral gaussian splatting. *arXiv preprint arXiv:2505.21890* (2025).
24. Wang, Z., Wu, S., Xie, W., Chen, M. & Prisacariu, V. A. Nerf-: Neural radiance fields without known camera parameters. *arXiv preprint arXiv:2102.07064* (2021).
25. Kruse, F. A. *et al.* The spectral image processing system (sips): Interactive visualization and analysis of imaging spectrometer data. In *Remote Sensing of Environment* (1993).
26. Kingma, D. P. & Ba, J. Adam: A method for stochastic optimization. In *International Conference on Learning Representations* (2015).
27. Xia, Y., Tang, H., Timofte, R. & Gool, L. V. Sinerf: Sinusoidal neural radiance fields for joint pose estimation and scene reconstruction (2022). [2210.04553](https://arxiv.org/abs/2210.04553).
28. Bian, W., Wang, Z., Li, K., Bian, J.-W. & Prisacariu, V. A. Nope-nerf: Optimising neural radiance field with no pose prior (2023). [2212.07388](https://arxiv.org/abs/2212.07388).
29. Schmidt, V.-M. *et al.* Handheld hyperspectral imaging as a tool for the post-mortem interval estimation of human skeletal remains. *Heliyon* **10**, e25844, DOI: [10.1016/j.heliyon.2024.e25844](https://doi.org/10.1016/j.heliyon.2024.e25844) (2024).
30. Knecht, S. *et al.* Sex estimation from long bones: a machine learning approach. *Int. J. Leg. Medicine* **137**, 1887–1895, DOI: [10.1007/s00414-023-03072-4](https://doi.org/10.1007/s00414-023-03072-4) (2023).
31. Youssef, D., Motto-Ros, V. & Abdel-Harith, M. Machine learning-based novel approach of image analysis in libs elemental mapping for discriminating archeological human bone. *Spectrochimica Acta Part B: At. Spectrosc.* **224**, 107108, DOI: [10.1016/j.sab.2024.107108](https://doi.org/10.1016/j.sab.2024.107108) (2025).
32. NIREOS. HERA VNIR Hyperspectral Camera. <https://nireos.com/product/hera-vnir/>. Accessed: 2026-05-28.
33. Mondal, A., Nguyen, C., Ma, X., Elbanna, A. E. & Carlson, J. M. Network models for characterization of trabecular bone. *arXiv preprint arXiv:1811.00092* (2018).
34. McGregor, M., Patel, S., McLachlin, S. & Vlasea, M. Architectural bone parameters and the relationship to titanium lattice design for powder bed fusion additive manufacturing. *Addit. Manuf.* **47**, 102273, DOI: [10.1016/j.addma.2021.102273](https://doi.org/10.1016/j.addma.2021.102273) (2021).
35. Chandramohan, D. *et al.* Bone material analogues for pet/mri phantoms. *arXiv preprint arXiv:1910.11315* (2019).
36. Peña Fernández, M. *et al.* Time-resolved in situ synchrotron-microct: 4d deformation of bone and bone analogues using digital volume correlation. *Acta Biomater.* **131**, 424–439, DOI: <https://doi.org/10.1016/j.actbio.2021.06.014> (2021).
37. Gao, K. *et al.* Neural radiance fields in 3d vision: A comprehensive review. *Comput. Vis. Media* 1–55, DOI: [10.26599/cvm.2026.9450535](https://doi.org/10.26599/cvm.2026.9450535) (2026).
38. Popescu, M.-C., Balas, V. E., Perescu-Popescu, L. & Mastorakis, N. Multilayer perceptron and neural networks. *WSEAS Trans. Cir. Sys.* **8**, 579–588 (2009).
39. Ku, K., Jubery, T. Z., Krishnamurthy, A. & Ganapathysubramanian, B. Hs-3d-nerf: 3d surface and hyperspectral reconstruction from stationary hyperspectral images using multi-channel nerfs. *arXiv preprint arXiv:2602.16950* (2026).
40. Cai, Y., Liu, X. & Cai, Z. BS-Nets: An end-to-end framework for band selection of hyperspectral image. *arXiv preprint arXiv:1904.08269* (2019).
41. Deb, D. & Verma, U. Correlation-based band selection for hyperspectral image classification. *arXiv preprint arXiv:2501.14338* (2025).
42. Hu, E. J. *et al.* Lora: Low-rank adaptation of large language models. *arXiv preprint arXiv:2106.09685* (2021).

## Acknowledgements

The authors gratefully acknowledge NIREOS for providing HERA hyperspectral camera<sup>32</sup> used for data collection in this study.

## Ethics declarations

Mouse knee joints from male one year old C57BL/6 were obtained from Queen Mary University of London. All procedures complied with the UK Animals (Scientific Procedures) Act of 1986 and were performed under a UK Home Office Project Licence in accordance with the Queen Mary University of London Ethical Policy on the use of Animals in Scientific Research and the ARRIVE guidelines. The study was approved by ethical review board at the Queen Mary University of London. Mice were housed under standard conditions with a 12 h light/dark cycle, and unlimited access to food and water.

## **Author contributions statement**

G.T. initialised concepts and directions. S.A. provided the mouse sample. G.T., T.N., and N.S. acquired the dataset. N.S. conducted the experiments and analysed the results. T.N. and G.T. provided critical updates and suggestions that significantly enhanced the scope and direction of the research. All authors N.S., T.N., S.A. and G.T. reviewed and approved the final manuscript.