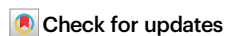


AI-Assisted segmentation and volumetric reconstruction of radiographs through multi-angular scintillation imaging

Received: 23 September 2025

Accepted: 9 May 2026

Published online: 04 June 2026



Seokho Kim^{1,11}, Byungjoon Bae^{1,11}, Do Wan Kim^{1,11}, Yoo Jin Lee², Siyeon Kim^{3,4}, Sungwoo Kim⁵, Jawoon Kim⁶, Cheong Beom Lee⁷, Yongmin Baek¹, Sree Sourav Das¹, Jiyoung Boo⁶, Jinho Choi⁸, Mona Zebbarjadi^{1,9}, Kyeounghak Kim¹✉, Sangeun Cho¹⁰✉, Dong Hyuk Park⁶✉ & Kyusang Lee^{1,9}✉

X-ray imaging serves as a fundamental tool for non-destructive inspection. Although conventional radiography is well suited for two-dimensional imaging, it cannot provide volumetric structure. Computed tomography provides three-dimensional reconstruction but remains constrained by bulky instrumentation, high radiation exposure, and cost. Here we demonstrate a patch-type scintillator integrated with multi-stage neural network that segments and reconstructs three-dimensional volumes from sparse angular two-dimensional radiographs. The scintillator is fabricated by electrospraying cellulose nanocrystals onto a bulk cellulose matrix, followed by dip-coating of perovskite, yielding a composite with enhanced radioluminescence under X-ray excitation. This flexible film conforms to complex geometries, enabling distortion-free and multi-angle imaging. Neural networks are trained on synthetic datasets and validated on experimentally acquired avian tibiotarsus radiographs, accurately reconstructing volumetric bone structures. This approach serves as a proof-of-concept for low-dose, accessible artificial intelligence-enabled three-dimensional X-ray imaging, demonstrating the feasibility of recovering macroscopic three-dimensional morphology from as few as three sparse projections.

X-ray imaging is a foundational technology for non-invasive inspection across medicine, security, and industrial evaluation, owing to its deep penetration and rapid structural visualization^{1–4}. In clinical practice, radiography affords rapid frontline assessment of medical conditions,

including fractures, air or fluid accumulation in the chest or abdomen, infection such as pneumonia, and tumors^{5–8}. However, conventional X-ray imaging is limited to two-dimensional (2D) projections, inevitably compressing three-dimensional (3D) spatial information into

¹Department of Electrical and Computer Engineering, University of Virginia, Charlottesville, USA. ²Department of Radiology, University of California San Francisco, San Francisco, USA. ³Department of Orthopedic Surgery, Samsung Medical Center, Sungkyunkwan University School of Medicine, Seoul, Republic of Korea. ⁴Department of Clinical Research Design and Evaluation, Samsung Advanced Institute for Health Sciences and Technology, Sungkyunkwan University, Seoul, Republic of Korea. ⁵Department of Radiation Oncology, Asan Medical Center, Seoul, Republic of Korea. ⁶Department of Chemical Engineering, Inha University, Incheon, Republic of Korea. ⁷Department of Chemical Engineering, Hanyang University, Seoul, Republic of Korea. ⁸Department of Materials Science and Engineering, University of Michigan, Ann Arbor, USA. ⁹Department of Materials Science and Engineering, University of Virginia, Charlottesville, USA. ¹⁰Division of System Semiconductor, Dongguk University, Seoul, Republic of Korea. ¹¹These authors contributed equally: Seokho Kim, Byungjoon Bae, Do Wan Kim. ✉e-mail: chemekim@hanyang.ac.kr; sangeun.c@dongguk.edu; donghyuk@inha.ac.kr; kyusang@virginia.edu

simple grayscale intensity maps, eliminating the critical depth cues. The dimensional reduction remains an intrinsic limitation when evaluating anatomically complex regions where structures overlap. For instance, breast lesions are notoriously difficult to precisely diagnose via standard chest X-ray mammography projections due to the complex 3D geometry of the breasts^{9,10}.

Computed tomography (CT) overcomes these limitations by providing high-fidelity volumetric reconstructions¹¹ of internal structures. However, its deployment in emergency, field, or point-of-care settings is restricted by bulky hardware and operational costs. Alternatively, digital tomosynthesis (DT) generates depth-resolved images by acquiring multiple projections over a limited angular range, providing sectional views at different depths utilized for mammography and chest imaging. Rotational angiography CT (C-Arm) provides 3D reconstructions through rotational imaging around the patient with more compact equipment, delivering real-time procedural guidance for interventional radiology and neurosurgery. However, these techniques still demand specialized equipment and a higher radiation dose than conventional radiography. Furthermore, precise segmentation of anatomical regions of interest (ROI), including bones, organs, lesions, tissues, and vasculature, is critical for diagnostic accuracy^{12–18}, yet has relied on labor-intensive manual annotation by expert clinicians.

Recent advancements in artificial intelligence (AI) technologies have significantly improved both the accuracy and efficiency of anatomical segmentation and 3D reconstruction from 2D X-ray images^{16–18}. Deep-learning frameworks leverage richly annotated datasets to train convolutional and transformer-based networks that infer volumetric structures from limited projections, thereby achieving 2D-to-3D reconstruction. Segmentation accuracy and reconstruction fidelity hinge critically on the quality and diversity of acquired images. However, conventional X-ray imaging systems, constrained by single-angle acquisition, introduce geometric distortions and occlusions that aggravate structural identification.

Addressing these challenges necessitates an interdisciplinary approach that integrates advanced X-ray imaging hardware with sophisticated reconstruction algorithms. On the hardware front, multi-angle X-ray imagers capture complementary projections that enrich the input data space for reconstruction models. In parallel, reconstruction algorithms incorporate statistical shape priors to regularize the solution and constrain it to anatomically plausible manifolds. Such innovations provide patient-specific 3D models, supporting improved diagnostic accuracy, deformity assessment, morphometry and surgical planning.

Here, we demonstrate a hardware and software co-designed approach that integrates a conformable perovskite-based scintillator with multi-stage neural networks to achieve high-precision 3D reconstruction from a limited set of 2D X-ray projections. Our flexible scintillator comprises Cs₃MnBr₅ perovskite uniformly nucleated within a mechanically flexible and thin cellulose matrix via an electrospray deposition of cellulose nanocrystals (CNs), followed by immersion in a mixed CsBr and MnBr₂ precursor. This approach yields a large-area and uniform film whose conformal application to complex specimen geometries eliminates parallax and distortion, thereby delivering sharp and multi-angle X-ray luminescence images. This lead-free, water-based clean-chemistry approach produces a non-toxic, conformable scintillator that supports dense perovskite dispersion at reduced thickness, enabling distortion-free multi-angle acquisition.

On the software side, each planar projection is first processed by the Segment Anything Model (SAM) algorithm¹⁹ to delineate target anatomical features. The segmented images are then independently processed using the TripoSR algorithm²⁰, which reconstructs an individual volumetric estimate from each view. A volumetric fusion neural network then consolidates these partial reconstructions into a single coherent 3D model, interpolating regions of low overlap of missing data. This integrated platform, featuring a flexible scintillator with

neural network-assisted segmentation and reconstruction algorithms, enables high-fidelity volumetric X-ray imaging from a limited number of projections, offering a pathway toward dose-efficient imaging with broad potential for point-of-care clinical and industrial non-destructive testing applications. (Supplementary Table 1).

Overview of Three dimensional Image using Flexible Platform

X-ray imaging is achieved by transmitting X-rays through an object, where variations in absorption generate contrast, revealing internal structures. We have developed a multi-angle X-ray imaging system that employs conformal patch-type scintillator film wrapped around the specimen, thereby enabling flexible selection of incident X-ray angle and distortion-free acquisition of multi-angle projections (Fig. 1a). The overall imaging and reconstruction pipeline is summarized in Supplementary Fig. 1. Upon X-ray excitation, the flexible film composed of Cs₃MnBr₅ perovskite dispersed within a cellulose matrix emits green luminescence (519 nm). With a thickness of only ~220 μm, this scintillator film conforms seamlessly to the shape of the target object. Such conformability maintains close contact between the specimen and the imaging plane, resulting in enhanced edge definition while minimizing distortion. This allows the scintillator film to be wrapped around curved surfaces, such as the human body (Fig. 1b). Furthermore, our approach decouples the scintillator from the detector, allowing the use of high-resolution imaging sensors to capture detailed X-ray projections suitable for tomographic algorithms.

For reconstructing 3D structures from multi-viewpoints of X-ray projections, a region-identifying neural network detects diagnostically relevant areas, then SAM produces precise and pixel-level segmentation across the entire dataset. Each segmented image is fed into TripoSR, a transformer-based network that generates a high-fidelity volumetric model of the specimen. The resulting partial 3D models are subsequently fused through a volumetric U-net-based fusion network, which reconciles inter-view inconsistencies and fills occluded regions to yield a coherent 3D model. This approach provides selective high-precision 3D reconstruction, particularly valuable for long-term monitoring of skeletal pathologies such as osteosarcoma, and holds promise for deployment in resource-limited clinical settings.

Result

Fabrication and analysis of flexible thin-film scintillator

Scintillators convert high-energy radiation, such as X-rays and gamma rays, into visible light; thus, they are essential for radiation detection and X-ray imaging. Among various scintillating materials, metal-halide perovskites have emerged as promising candidates due to their high radioluminescence (RL) quantum yield and tunable emission spectra. In this work, we employed a lead-free, all-inorganic metal-halide perovskite Cs₃MnBr₅ as the active scintillation medium, which exhibits intense green luminescence under X-ray excitation. However, achieving a homogeneous perovskite film over large areas poses significant fabrication challenges. Furthermore, conventional perovskite-based scintillator films usually require hundreds of micrometers of thickness to effectively absorb transmitted X-rays, leading to substantial material consumption and cost inefficiencies^{21,22}. To address this limitation while maintaining structural integrity and absorption efficiency, we have fabricated a composite scintillator in which Cs₃MnBr₅ perovskite is embedded within a 220 μm thick cellulose fiber matrix. This hybrid architecture permits low-cost and scalable fabrication of a flexible scintillator film capable of X-ray imaging.

Figure 2a presents a schematic illustration of the scintillator fabrication process. Achieving a uniform Cs₃MnBr₅ coating on bundled cellulose fibers via solution process is inherently challenging due to the coffee-ring effect. The capillary flows during drying drive solute migration toward the fiber edges, resulting in a physical thickening of the crystal layer at the perimeter. This mass transport leads to a ring-like deposition pattern where the intense edge emission arises from

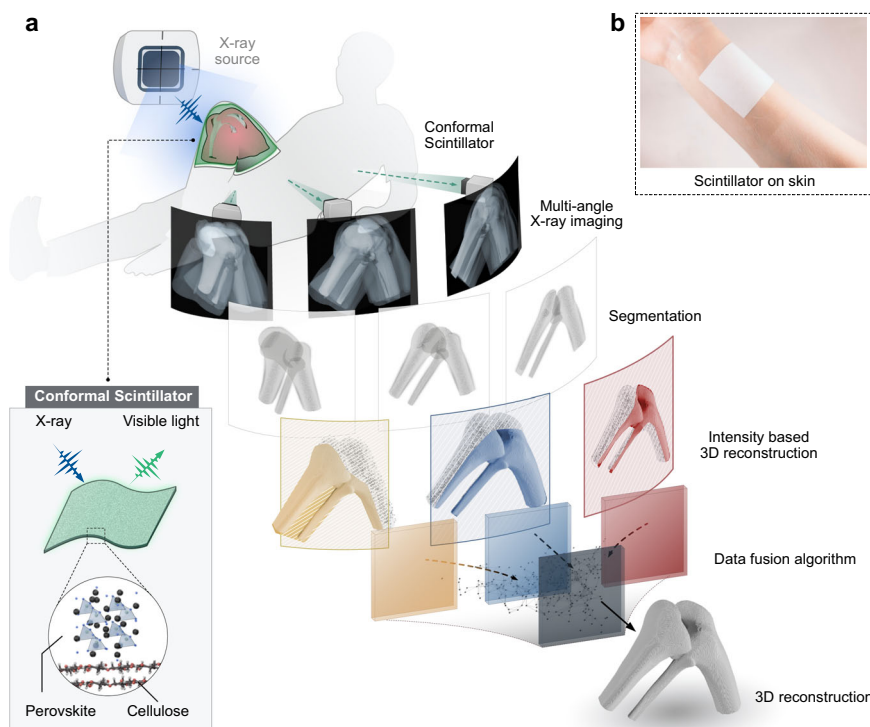


Fig. 1 | Workflow for X-ray image segmentation and 3D reconstruction via a flexible scintillator. **a** Schematic of the end-to-end imaging and reconstruction process. A flexible scintillator adheres tightly to the target object through a transparent adhesive patch, converts incident X-rays into a visible light. Sequential radiographs are captured from multiple angles, and each projection is processed by a segmentation algorithm to delineate the region of interest (ROI) and extract key

anatomical features. These segmented 2D images are then fed into a neural network, which volumetrically fuses the segmented data to generate a 3D reconstruction. **b** Photograph of the fabricated scintillator film conformally wrapped around a human forearm using a transparent dressing, illustrating its mechanical compliance and intimate contact with complex surface geometries.

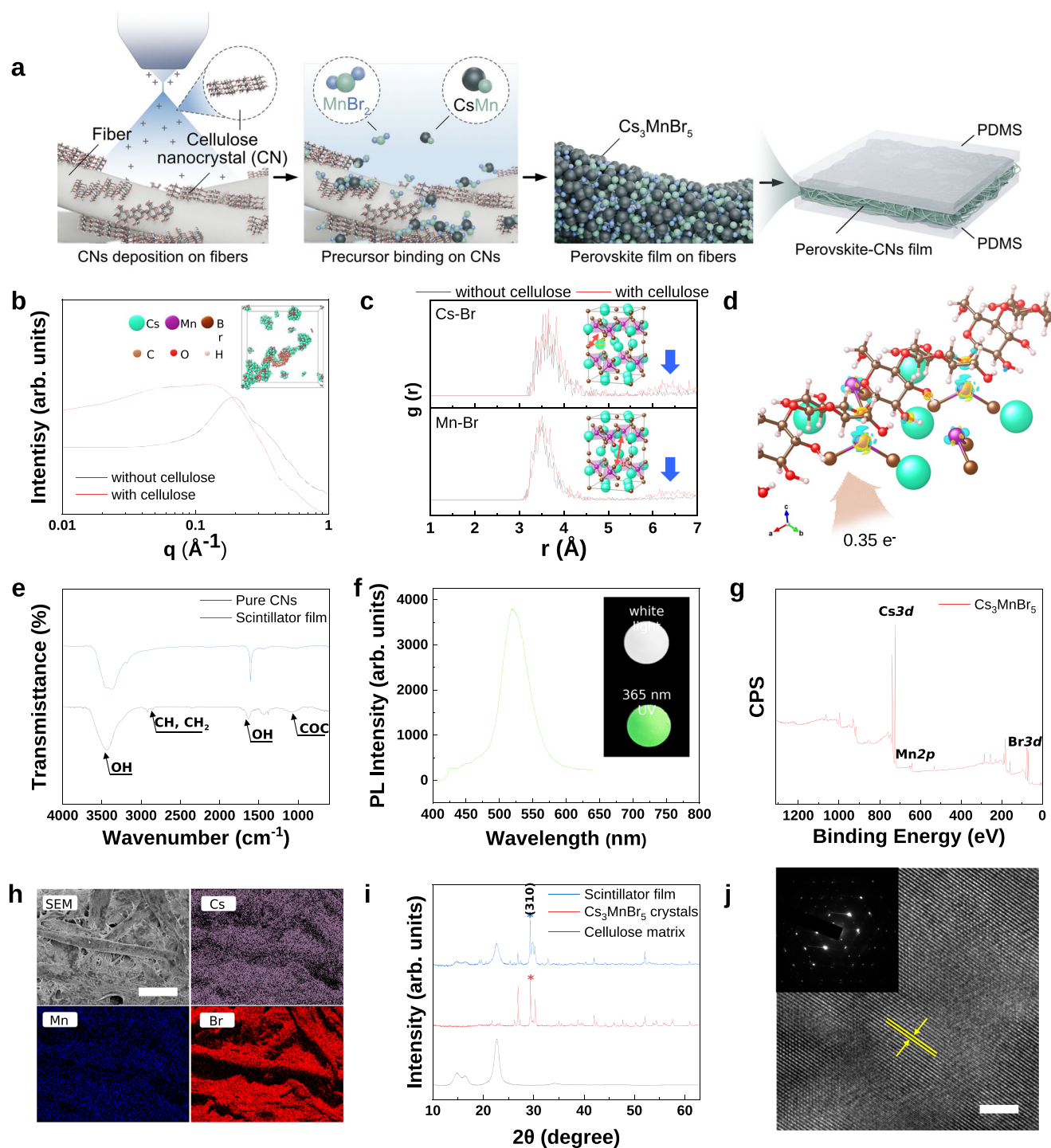
the accumulation of emitting material. To mitigate this, we first applied CNs via electrospray onto the fiber matrix, establishing a high density of nucleation sites that disrupt the outward capillary flow and promote uniform distribution of the precursors (Supplementary Fig. 2). Subsequent annealing at 80 °C for 4 h strengthens hydrogen bonds between the CNs and the underlying cellulose matrix, locking in the interfacial architecture.

Then, we employed a dip-coating process that forms Cs_3MnBr_5 perovskite crystals uniformly across a high-surface-area cellulose scaffold. Conventional syntheses of Cs_3MnBr_5 perovskite crystals, including HBr solution-based coating²³, lyophilization-assisted recrystallization (LARC)^{24,25}, hot-injection^{26,27}, and mechano-solid-state reactions methods²⁸, often rely on hazardous solvents and multistep procedures. In contrast, our approach uses only deionized (DI) water and mild annealing (50 °C–70 °C), significantly reducing chemical complexity and environmental impact. In our process, CNs-coated cellulose matrix was immersed in DI water containing 3 mmol CsBr and 1 mmol MnBr_2 precursors, allowing sufficient interaction between the β -D-glucose monomers in cellulose and the dissolved precursors. The nucleation of perovskite nanocrystals is driven by Lewis acid–base coordination, where hydroxyl (–OH) and glycosidic (C–O–C) groups in cellulose serve as electron donors to the vacant d orbitals of heavy metal ions in the precursors. This interaction fosters strong electrostatic binding with the precursors surrounding the cellulose scaffold (Supplementary Fig. 3a), resulting in the formation of stable coordinate covalent bonds that anchor Cs_3MnBr_5 nanocrystals with a grain size of 10–20 nm along the fiber network (Supplementary Fig. 3b). The fabrication process exhibits excellent repeatability, as demonstrated by structural and compositional consistency across five independently fabricated batches evaluated by non-resonance Raman spectroscopy (Supplementary

Fig. 4). The Raman peaks of the Cs_3MnBr_5 at 111.5 cm^{-1} , 160.5 cm^{-1} , and 167.8 cm^{-1} were reproducibly observed across all films, and quantitative comparison of baseline-corrected peak intensities yields a relative standard deviation (RSD) of 4.2%, confirming high batch-to-batch uniformity.

Following the dip-coating process, the composite films were dried in a vacuum oven for 18 h at a temperature range of 50 °C–70 °C, allowing gradual evaporation of the solvent while preserving the crystallized perovskite on the network of cellulose fibers. Given the inherent water solubility of Cs_3MnBr_5 perovskite crystals, moisture exposure poses a severe degradation risk, necessitating protective encapsulation. To address this, we applied and cured a thin transparent polydimethylsiloxane (PDMS) overlayer on the composite film, imparting superior environmental resilience to scintillators. PDMS encapsulation mitigates moisture-induced ion migration and phase degradation under high-humidity conditions (Supplementary Fig. 5)^{29,30}. To evaluate signal stability under practical imaging conditions, RL image measurements were initiated following a 30-minute post-fabrication stabilization period and recorded at 20 min intervals over 80 min, a duration exceeding the typical multi-angle projection acquisition time. Total intensity drift over this period was 1.3%, confirming stable RL emission throughout a realistic imaging workflow. Moreover, the PDMS layer minimized cumulative mechanical damage of cellulose due to repeated bending by reducing localized stress concentrations (Supplementary Figs. 6 and 7). Therefore, the PDMS encapsulation enables the flexible perovskite–cellulose scintillators to maintain structural integrity and optical performance under realistic operating conditions.

To investigate the effect of CNs on the crystallization of Cs_3MnBr_5 , molecular dynamics (MD) simulations were performed. In the absence of CNs, CsBr and MnBr_2 precursors tended to aggregate randomly



(Supplementary Fig. 8a). However, when CNs were present, precursors preferentially nucleated and crystallized around the CNs (Inset of Fig. 2b). Complementary small-angle X-ray scattering (SAXS) profiles extracted from the MD simulations also showed the broad SAXS feature at lower q values in the precursors with CNs, implying enhanced particle coalescence with effective crystallization³¹ (Fig. 2b). Analysis of the radial distribution functions (RDFs) for Cs-Br and Mn-Br pairs further substantiated this effect. Secondary peaks corresponding to bulk Cs_3MnBr_5 (5.35 Å and 6.59 Å for Cs-Br; 6.20 Å and 6.55 Å for Mn-Br, Supplementary Fig. 9) became significantly more intense in the CNs-containing system (Fig. 2c). Furthermore, the velocity autocorrelation functions (VACFs) decayed more rapidly in the presence of CNs compared to the system without CNs, indicating reduced ionic

mobility and accelerated crystallization kinetics of Cs_3MnBr_5 when CNs are incorporated^{32,33} (Supplementary Fig. 10).

Density functional theory (DFT) calculations were carried out to further investigate the thermodynamic favorability of Cs_3MnBr_5 crystallization with CNs. The simulated formation energy (E_f) of Cs_3MnBr_5 when interfaced with CNs (2.46 eV) was substantially lower than that in the absence of CNs (3.76 eV), indicating that CNs render nucleation of Cs_3MnBr_5 perovskite more thermodynamically favorable. Charge density difference and Bader charge analyses further revealed a charge transfer from the Mn site to the O atoms in cellulose, indicating strong Mn–O interactions at the interface (Fig. 2d). The directional electron redistribution is illustrated in the figure by an arrow pointing from Mn to O, accompanied by the transferred charge, highlighting the

Fig. 2 | Fabrication and Characterization of Cellulose-Perovskite hybrid Scintillator Films. **a** Schematic of the flexible scintillator fabrication via electro-spraying of cellulose nanocrystals (CNs) followed by immersion in deionized water with CsBr/MnBr₂ precursors to induce Cs₃MnBr₅ crystallization. A poly-dimethylsiloxane overcoat ensures environmental stability. **b** Simulated small-angle X-ray scattering patterns. Broad lower q features with cellulose indicate enhanced precursor aggregation. Inset: Molecular dynamics simulation of the precursor structure. **c** Radial distribution function $g(r)$ for Cs-Br and Mn-Br pairs. Secondary correlation peaks (blue arrows) correspond to interatomic separations of 6.59 Å and 6.55 Å. Inset: Atomic site within the Cs₃MnBr₅ unit cell. Red arrows indicate the corresponding atoms. **d** Density functional theory-derived charge density difference and Bader charge transfer from Mn to O atoms, indicating interfacial stabilization of perovskite nucleation. **e** Fourier transform infrared of CNs and the composite. Retention of hydroxyl stretching (3444 cm⁻¹) and bending (1631 cm⁻¹) mode alongside -CH or -CH₂ stretching (2850–2950 cm⁻¹) and glycosidic bond (C–O–C, 1068 cm⁻¹) bands, confirms chemical integrity. **f** Photoluminescence

spectrum (λ_{ex} = 365 nm), showing green emission at 519 nm. Inset: photographs of the scintillator film under white light and UV illumination. **g** X-ray photoelectron spectroscopy survey confirming elemental composition via binding energy peaks for Cs 3d (724.02 eV), Mn 2p (641.05 eV), and Br 3d (68.86 eV). **h** Scanning electron microscopy image and corresponding energy dispersive spectroscopy maps for Cs, Mn, and Br, demonstrating homogeneous dispersion. The image shown is representative of 5 independent batches, with similar results. Scale bar, 50 μ m. **i** X-ray diffraction analysis of the scintillator film, reference Cs₃MnBr₅ powder and pure cellulose. Peaks at 26° and 29° correspond to the (213) and (310) planes of tetragonal Cs₃MnBr₅; additional diffraction peaks at 14.8°, 16.4°, and 22.7° are observed in both cellulose and scintillator films. **j** High-resolution transmission electron microscopy of a Cs₃MnBr₅ crystallite on cellulose, showing (310) lattice fringes with an interplanar spacing of a 3.0 Å. Inset: The selected area electron diffraction pattern with arrows indicating the (310) plane. Yellow arrows indicate a lattice spacing (gap) of 3 Å. Similar results were obtained in 4 independent batches. Scale bar, 5 nm.

electrostatic nature of the interfacial interaction. This charge redistribution originates from the difference in electronegativity between Mn and O, where the more electronegative oxygen atoms attract electron density from the Mn centers. As a result, partial charges are induced, giving rise to electrostatic attraction at the interface, which supports the presence of strong interfacial interactions between the precursors and the cellulose matrix. Together with our MD simulation results, these first-principles insights demonstrate that the cellulose scaffold actively facilitates the crystallization of Cs₃MnBr₅, primarily through the interactions between Mn and O atoms in perovskite and cellulose, respectively.

To experimentally investigate the chemical interactions between CNs-coated cellulose matrix and the incorporated Cs₃MnBr₅ perovskite crystals, we conducted Fourier transform infrared (FT-IR) spectroscopy on both the unmodified cellulose matrix and the perovskite-cellulose hybrid (Fig. 2e). In the pristine cellulose spectrum³⁴, the β -D-glucose monomer exhibits a hydroxyl (O–H) stretching vibration at 3444 cm⁻¹ and a corresponding bending mode at 1631 cm⁻¹. Upon perovskite incorporation, distinct stretching vibrations³⁴ associated with –CH and –CH₂ moieties are observed between 2850 and 2923 cm⁻¹. Furthermore, the glycosidic linkage (C–O–C), which originates from the β -1,4-glycosidic bonds between β -D-glucose units and forms the backbone of the cellulose polymer³⁵, is evidenced by a prominent absorption band at 1068 cm⁻¹. These bands undergo subtle shifts and intensity changes, confirming the establishment of hydrogen-bond and coordinate-bond interactions at the cellulose and perovskite interface.

Furthermore, the CNs with a net negative charge through their functional groups were quantitatively verified through zeta potential measurements (Supplementary Fig. 11). The negatively charged surface of CNs facilitates electrostatic interactions with Cs⁺ and Mn²⁺ ions in DI water, potentially leading to their localized accumulation near the CNs surface^{36–38}. The subsequent interaction of these ions with Br⁻ in the solution could lead to a locally elevated concentration of precursors, which promotes nucleation under suitable conditions. Thereafter, the adsorption and association of additional ions onto these initial sites on the cellulose matrix contribute to further crystal growth. Figure 2f shows the characteristic green emission at 519 nm from the Cs₃MnBr₅ crystal under 365 nm UV excitation, corresponding to an external quantum efficiency (EQE) of 46.89% as determined by quantitative integrating sphere measurements (Supplementary Fig. 12). This emission originates from a radiative transition of the excited electron from the ⁴T₁(G) state to the ground state ⁶A₁ within the MnBr₄²⁻ tetrahedral units, which serve as the luminescent centers. The transition corresponds to an intra-3d electronic transition of Mn²⁺ ions, typically resulting in a narrow-band, sharp emission. The inset images are photographs of the scintillator film under white and a 365 nm UV

light, respectively, showing its white appearance in the absence of excitation and bright green emission under UV illumination^{39,40}.

To confirm the chemical integrity of the synthesized material, X-ray photoelectron spectroscopy (XPS) was conducted on fully dried scintillator films (Fig. 2g). The XPS spectra showed binding energies corresponding to Cs 3d_{5/2} (724.02 eV), Mn 2p_{3/2} (641.05 eV), and Br 3d_{5/2} (68.86 eV), confirming the successful formation of the intended Cs-Mn-Br chemical bonds. Quantitative elemental analysis further indicated atomic composition of Cs (35.07%), Mn (10.66%), and Br (54.27%), in close agreement with the stoichiometry of Cs₃MnBr₅ perovskite crystals. These results validate that the self-assembled precursors have crystallized into the targeted perovskite composition with preserved stable chemical bonding within the cellulose matrix, suitable for scintillator applications.

Energy-dispersive X-ray spectroscopy (EDS) was conducted to assess the spatial uniformity of Cs₃MnBr₅ perovskite within the cellulose scaffold (Fig. 2h). The scanning electron microscopy (SEM) image (the top left panel) shows the cellulose matrix consisting of intricately intertwined fibers with diameters ranging from 2 μ m to 20 μ m. Elemental maps for Cs, Br, and Mn, displayed in a clockwise order from the top right, exhibit a homogeneous coating of perovskite species along the fiber surfaces. Quantitative EDS analysis shows atomic ratios of Cs, Mn, and Br at 38.6, 10.2, and 51.1%, respectively (Supplementary Fig. 13), closely aligning with the targeted stoichiometry and further confirmed with complementary XPS measurements. This concordance confirms the uniform dispersion of the scintillator components throughout the cellulose framework.

Following confirmation of its elemental composition, we verified the crystallographic phase of the Cs₃MnBr₅ scintillator film by X-ray diffraction (XRD) analysis (Fig. 2i). Cs₃MnBr₅ crystallizes in a tetragonal crystal structure, wherein Mn²⁺ centers are coordinated by four Br⁻ anions, and Cs⁺ ions occupy the interstitial sites. The lattice parameters are $a = b = 9.71$ Å, $c = 15.46$ Å, with orthogonal angles of $\alpha = \beta = \gamma = 90.00^\circ$ (Supplementary Fig. 14). The XRD spectrum of the composite film (blue line) exhibits distinct peaks attributable to both the CNs (black line) and the Cs₃MnBr₅ phase (red line), indicating the coexistence of both components with preserved structural integrity. The diffraction peak at 29.8° (denoted by asterisks) corresponding to the (310) crystal plane with a d-spacing of 3.0 Å was enhanced while the (213) plane diffraction at 26.4° was diminished relative to the bulk Cs₃MnBr₅ perovskite reference (Supplementary Fig. 14a), indicating a preferential orientation or slight lattice distortion of [MnBr₄]²⁻ tetrahedra, likely induced by Mn–O interactions with the CNs, as well as by the accumulation of cesium bromide on the surface. The characteristic XRD peaks observed at 14.8°, 16.4°, and 22.7° align well with the (110), (110), and (200) planes of CNs, respectively.

High-resolution transmission electron microscopy (HR-TEM) analyses further confirm the formation of a well-ordered Cs₃MnBr₅

crystal on the cellulose matrix (Fig. 2j). Distinct lattice fringes with a periodic spacing of 3.0 Å are clearly resolved, matching with (310) plane spacing of the tetragonal perovskite. The selected area electron diffraction (SAED) pattern in the inset of Fig. 2j, along with the fast Fourier-transformed (FFT) image (Supplementary Fig. 14b), reveals distinct diffraction spots, indicating the uniformly oriented crystalline domains. In contrast, the control sample without CNs (Supplementary Fig. 15) exhibits disordered lattice fringes and diffuse diffraction patterns, indicating significantly lower crystallinity. These observations validate the high crystallinity of Cs₃MnBr₅ on the cellulose matrix.

Radioluminescence characteristics of scintillator

Halide perovskite-based scintillators that incorporate heavy metal cations and halogen anions demonstrate high X-ray absorption and efficient radiative recombination, attributed to the quantum confinement effect and extended carrier lifetimes^{41–45}. Conventional perovskite scintillators, however, typically require a film thickness up to 800 µm for optimal scintillation performance, which presents significant fabrication challenges²². By contrast, our composite architecture utilizes a heterostructure wherein only a sub-micron-thick perovskite crystal layer is grown on a cellulose matrix pre-coated with CNs, ensuring effective scintillation performance with a total thickness of approximately 220 µm (Supplementary Fig. 16). The scintillation process within this composite system is further facilitated by X-ray absorption and charge transfer from the cellulose matrix, the dominant volumetric component, to the perovskite crystallites.

Figure 3a schematically illustrates the scintillation process within a heterostructure composed of Cs₃MnBr₅ and cellulose with energy bandgaps of 2.4 eV and 5.5 eV, respectively. In Cs₃MnBr₅, the radiative properties are primarily governed by the Mn²⁺ *d*–*d* transitions within a tetrahedral crystal field, which splits the Mn *d*-orbitals into discrete energy levels of ⁴E_g(D), ⁴T₂(D), ⁴A₁, ⁴E(G), and ⁴T₂(G)^{46,47}. An extended Mn–Mn distance of approximately 7 Å, derived from theoretical calculations, in Cs₃MnBr₅ leads to strong exciton confinement and efficient self-trapped–exciton emission^{25,48}. Consistent with this band alignment, UV-Vis absorption measurements yield optical band gaps of 2.02 eV for Cs₃MnBr₅ and 5.13 eV for cellulose (Supplementary Fig. 17), in reasonable agreement with theoretical predictions. Upon X-ray irradiation, both the perovskite and cellulose absorb high-energy photons via photoelectric effect, generating energetic electrons and excitons with absorption coefficients in X-ray spectrum on the order of 10⁵ cm^{−1} and 10³ cm^{−1}, respectively^{49,50}.

First-principles calculations were performed to probe the electronic structure of the cellulose–Cs₃MnBr₅ heterostructure. The (001) crystallographic plane of Cs₃MnBr₅ was selected as the interfacial surface for the heterojunction model, as it is the most thermodynamically stable and prevalent facet during crystal growth^{51–55}. The projected density of states (PDOS), referenced to the Fermi level, is presented in Fig. 3b. The PDOS of cellulose reveals that its lowest unoccupied molecular orbital (LUMO) lies at approximately 4.1 eV, whereas the Mn-3*d* states in Cs₃MnBr₅ are located at ~2.14 eV (Supplementary Fig. 18). To assess interfacial charge redistribution, we calculated work functions for the Br-terminated Cs₃MnBr₅ surface, cellulose, and the combined heterostructure (Supplementary Fig. 19). The intermediate work function of the heterostructure implies electrons diffuse from cellulose to Cs₃MnBr₅ upon contact, supported by the charge-density analysis in Fig. 3c. The three-dimensional charge density map (left) shows regions of depletion (blue) and accumulation (red), while the planar-averaged profile (right) highlights electron depletion in the cellulose region (blue shaded area) and accumulation in the Cs₃MnBr₅ layer (red shaded area). Integrating the charge difference yields a net transfer of electrons from cellulose to the perovskite, indicating a thermodynamically favorable donor–acceptor interaction at the interface (Supplementary Materials and Methods: Density Functional Theory Calculations). To further support the

interfacial charge transfer predicted by first-principles calculations, we performed high-resolution XPS measurements (Supplementary Fig. 20). The Cs 3*d*, Mn 2*p*, and Br 3*d* core levels in the hybrid film exhibit a consistent negative shift of approximately 0.5–0.6 eV relative to pristine Cs₃MnBr₅. This shift indicates enhanced electronic screening associated with increased electron density in the Cs₃MnBr₅ phase, in agreement with the DFT-predictions of electron transfer from cellulose to the perovskite.

Following absorption, excited carriers in the perovskite rapidly undergo thermalization via phonon interactions, losing excess energy until they relax to the lowest excited state (⁴T₁(G)). Meanwhile, high-energy excitons are induced in the cellulose matrix upon X-ray excitation and are subsequently transferred to the perovskite via interfacial charge transfer. The LUMO energy level of the cellulose is significantly higher than that of the ⁴E_g(D) level of Mn²⁺, facilitating efficient charge transfer (Fig. 3b). Subsequently, the well-defined *d*–*d* crystal field transitions in Mn²⁺, from ⁴T₁(G) to the ground state (⁶A₁)^{25,56}, enable efficient radiative recombination, resulting in the emission of 519 nm photons.

Figure 3d shows the RL spectra of the perovskite–cellulose hybrid film measured at various accelerating voltages. In all conditions, the composite exhibits a dominant emission peak at 519 nm, and the integrated RL intensity increases linearly with dose rate, indicating suitability for contrast-based X-ray imaging (summarized in Fig. 3e). The emission wavelength and full width at half maximum (FWHM) remain unchanged across dose rates, confirming that incorporation of the cellulose matrix enhances RL intensity through charge transfer without perturbing the intrinsic spectral signature of Cs₃MnBr₅.

To benchmark scintillation performance under X-ray excitation, we compared the RL output of our hybrid film against a commercial CsI:TI scintillator and two reported Mn-based scintillators under standardized conditions (200 kVp, 20 mA, 10 s) (Supplementary Fig. 21). The hybrid film exhibited a maximum RL intensity of 0.14 a.u. with a narrow FWHM of 45 nm. By comparison, the commercial CsI:TI reference showed a lower peak intensity of 0.06 a.u. with a substantially broader FWHM of 134 nm, while (Ph₄P)₂MnBr₄ and (Ph₃PH)₂MnBr₄ yielded intensities of 0.07 a.u. and 0.02 a.u. (FWHM of ~40 nm each). To provide a quantitative metric, we calculated light yield by calibrating integrated RL intensity against the CsI:TI standard (~54,000 photons MeV^{−1})^{57,58}. The hybrid film reached approximately 41,000 photons MeV^{−1}, 76% of the commercial benchmark, while the two Mn-based references yielded ~4400 and ~18,100 photons MeV^{−1}, respectively. Given that the composite is fabricated via a simplified DI water-based synthesis, this light yield confirms the competitive scintillation efficiency at substantially reduced processing complexity.

To assess the effect of CNs on scintillation performance, cellulose matrices with and without CNs were fabricated, and their RL intensities were compared over a controlled range of perovskite loadings (Fig. 3f). The perovskite content was precisely adjusted by dispensing fixed-concentration solution in volumes of 50 to 300 µL. At lower volumes (50–100 µL), the presence of CNs had minimal influence on RL intensity. In contrast, at volumes exceeding 150 µL, a pronounced enhancement was observed, with the 300 µL CNs-containing sample exhibiting approximately twice the RL intensity of its CNs-free counterpart. Linear fits of RL intensity versus solution volume yielded slopes of 4.6 (with CNs) and 1.3 (without CNs), indicating a ~3.5-fold increase in sensitivity. This enhancement arises from interfacial charge transfer facilitated by CNs, which accelerates radiative relaxation of X-ray-induced excitations, as well as uniform crystal growth across the surface promoted by CNs.

To probe the underlying photophysics, time-resolved photoluminescence (TR-PL) measurements were performed for perovskite–cellulose composites with varying Cs₃MnBr₅ weight percentages of 90, 75, 25, and 10 wt% (Fig. 3g). In the solution-processed

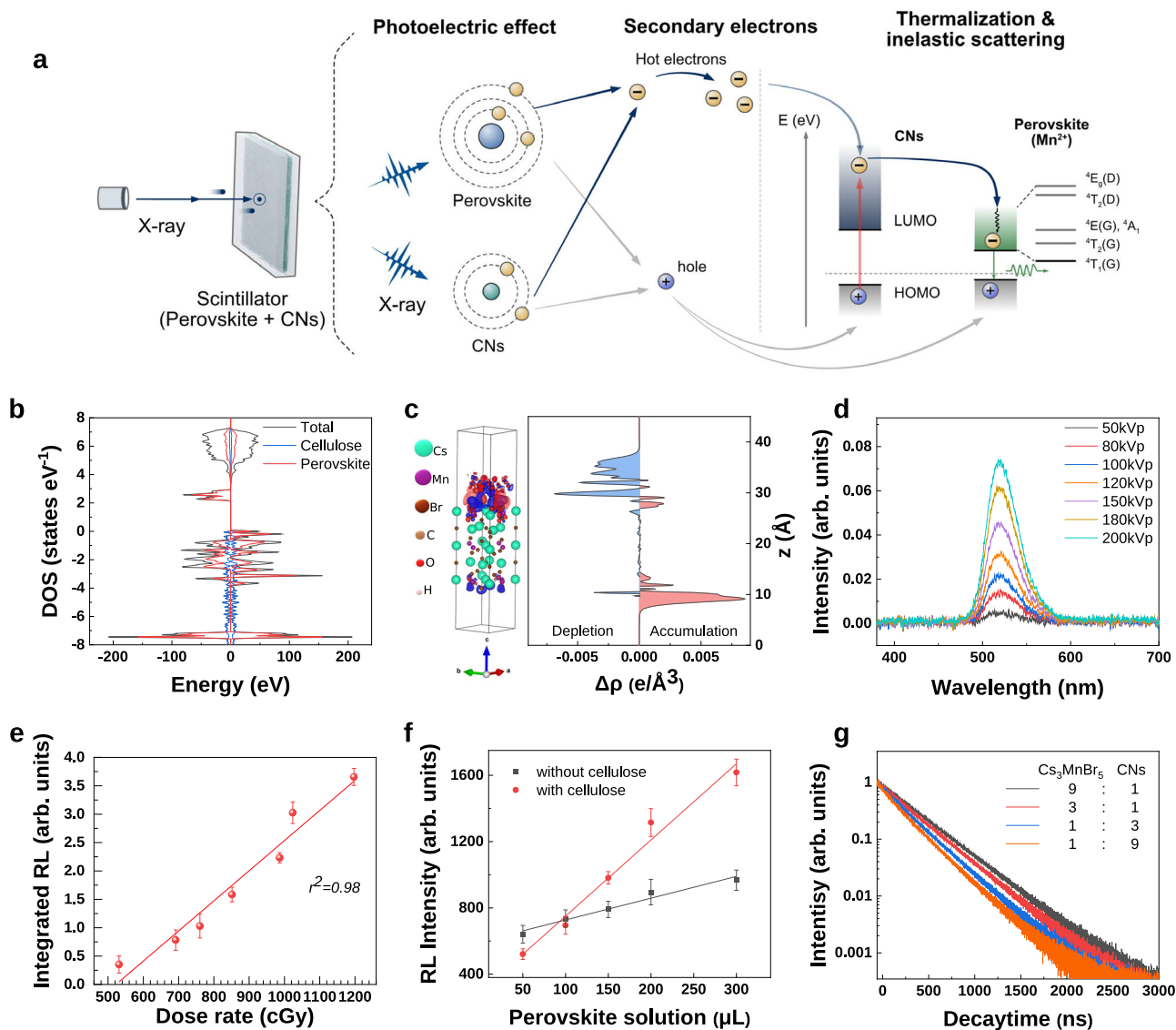


Fig. 3 | Radioluminescence (RL) of the Flexible Scintillator. **a** Schematic of the scintillation mechanism within the cellulose-perovskite hybrid composite. Upon X-ray irradiation, high-energy photons are absorbed in both Cs_3MnBr_5 and cellulose through the photoelectric effect, generating excitons. The high-energy excitons in cellulose transfer from the cellulose to the Cs_3MnBr_5 . Radiative recombination then proceeds through Mn^{2+} d–d transitions, yielding green RL emission. **b** First principle calculations of the density of states (DOS) for the cellulose-perovskite hybrid composite, showing the total (black), Cs_3MnBr_5 (red), and cellulose (blue) contributions relative to the Fermi level ($E - E_F$). The alignment reveals the relatively higher position of the cellulose LUMO compared to the Mn^{2+} $4E_g(D)$ levels in Cs_3MnBr_5 . **c** Iso-surface maps of the charge density difference (left), and the calculated charge depletion (blue) and accumulation (red) regions with the corresponding net charge transfer for the cellulose-perovskite hybrid composite (right). **d** RL spectra of the perovskite–cellulose hybrid film, measured under varying excitation voltages and beam currents, reveal a consistent emission peak at 519 nm across all excitation conditions. The RL intensity increases linearly with the

incident dose rate, confirming its applicability for contrast-based X-ray imaging. **e** The plot of RL intensity versus X-ray dose rate demonstrates a linear response, supporting the suitability of film for quantitative imaging applications. Data are presented as mean values $\pm$ standard deviation ($n = 5$ independent batches). **f** RL response of cellulose matrices with and without cellulose nanocrystals (CNs) measured as a function of deposited fixed-concentration perovskite solution volume (50–300 μL). CNs produce minimal change in RL at $\leq 100 \mu\text{L}$, while it causes a pronounced enhancement at larger volumes, yielding nearly double the RL intensity at 300 μL and approximately 3.5-fold steeper slope in RL–volume linear fits. Data are presented as mean values $\pm$ standard deviation ($n = 5$ independent batches for each condition). **g** Time-resolved photoluminescence (TR-PL) of perovskite-cellulose composites with varying Cs_3MnBr_5 weight percentages (90, 75, 25, and 10 wt%). Reducing the perovskite fraction shortens the PL decay constant, consistent with accelerated radiative recombination by dynamics in lower-perovskite-content films.

composites, the measured PL lifetimes decrease monotonically from 355.35 μs , 324.24 μs , 287.35 μs , to 261.82 μs , respectively, as the perovskite fraction is reduced, consistent with accelerated radiative recombination dynamics in composites with lower perovskite content^{59,60}. This acceleration trend is most pronounced in the fully integrated hybrid scintillator film composed of cellulose paper, CNs, and Cs_3MnBr_5 , which exhibits the shortest lifetime of 247.61 μs (Supplementary Fig. 22).

X-ray Imaging using a flexible thin-film scintillator

Conventional flat-panel X-ray detectors, which integrate rigid photodetectors on a TFT array with bulk scintillator film, are prone to geometric distortion and limited spatial accuracy when imaging non-planar structures. Although such detectors generally eliminate the veiling glare and uneven field response characteristic of image intensifiers, their planar architecture constrains close object coupling and can introduce parallax errors at oblique incidence. By

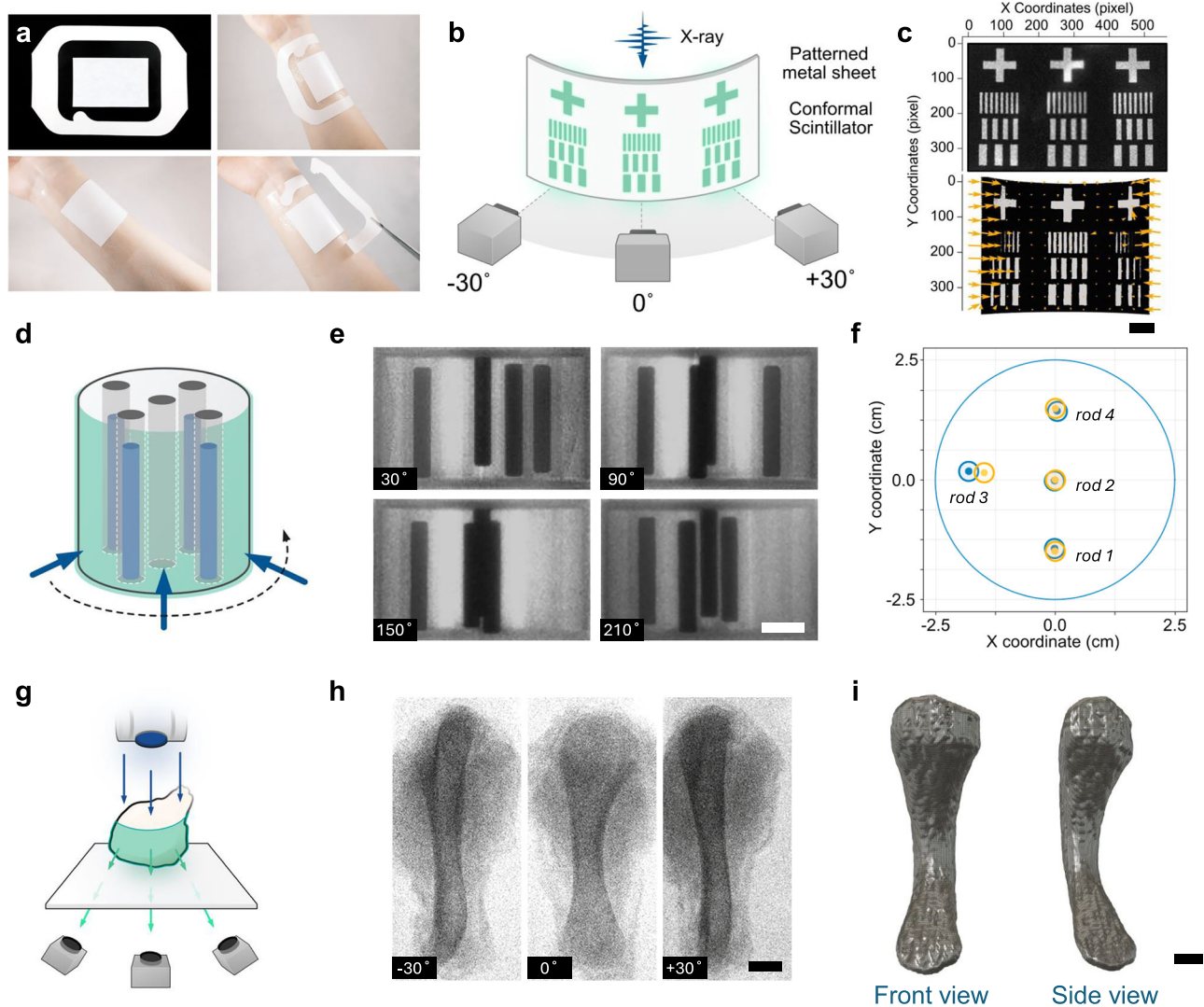


Fig. 4 | Conformal multiangle X-ray imaging with a flexible scintillator.

a Sequential application of the flexible scintillator patch (i) the film is bonded to an adhesive backing, (ii) the assembly is affixed to the wrist with a protective guard, (iii) the guard is removed, and (iv) the scintillator remains conformally attached.

b Schematic illustration of the experimental setup for distortion-free imaging validation using a cylindrical calibration phantom with bar patterns. Individual projections were acquired from multiple viewing orientations to evaluate imaging fidelity on curved surfaces. **c** Registration confirmed that the multi-angle dataset accurately reproduced the true spatial dimensions of the test patterns, in contrast to the geometric distortions observed in single-view acquisition. Scale bar, 10 mm. **d** Volumetric reconstruction of the four-pillar phantom, combining positional data from successive angular projections, enables precise pillar localization in three dimensions. **e** Multi-angle imaging of a cylindrical object, followed by 3D

recognition, reveals internal features obscured in a single projection view. The spatial arrangement and count of internal elements become unambiguous when visualized from multiple perspectives. Scale bar, 10 mm. **f** Comparison of the ground-truth rod positions (yellow) with the reconstructions derived from sparse multi-angle projections (blue) demonstrates that accurate recognition of the cylindrical object's internal geometry is achievable from a limited set of view.

g Experimental configuration for multi-angle X-ray imaging of a turkey tibiotarsus. **h** Fracture visualization in the tibiotarsus. The transverse crack appears as 0.4 cm in the frontal view (0°), 0.2 cm at -30° and 0.5 cm at +30°, illustrating the angular dependence of apparent fracture details. Scale bar, 10 mm. **i** 3D reconstructed result of observed ornithological leg. The left shows a front view of the reconstructed image, while the right shows a side view of it. Scale bar, 10 mm.

contrast, our approach employs a patch-type scintillator in conjunction with a high-resolution image sensor. As demonstrated in Fig. 4a, the flexible scintillator films conform seamlessly to the surface of target objects, thereby eliminating planar-detector distortions and improving edge definition. A transparent pressure-sensitive adhesive (PSA) ensures secure, wrinkle-free adhesion, enabling the capture of undistorted X-ray projections from multiple viewpoints without loss of spatial resolution.

To assess the flexibility and mechanical stability of the platform, we conducted a tensile stress-strain test using a universal testing machine (UTM) and performed a cyclic bending test on the scintillator film. The uniformly coated PDMS acts in parallel with the cellulose matrix in response to mechanical stress, significantly enhancing

mechanical robustness. Specifically, the tensile stress at maximum load increased from 10.15 MPa to 16.56 MPa, and the elastic modulus improved from 655.45 MPa to 1355.41 MPa, representing an approximately two-fold increase (Supplementary Fig. 6). Bending induces alternating tensile and compressive strains at the perovskite on cellulose composite, which can initiate microcracks and degrade luminescence performance in folded regions. To quantitatively analyze the effect, we subjected 20 μm -thick films to 1000 bending cycles at a radius of 2 cm. PL measurements taken after repeated bending and relaxation revealed that the scintillator film without PDMS exhibited structural fracture after approximately 270 bending cycles, accompanied by a sharp decrease in PL intensity at the folded region. In contrast, the PDMS-coated scintillator film showed no visible damage

even after 1000 cycles, indicating that the composite architecture maintains structural integrity and emission efficiency under prolonged mechanical deformation (Supplementary Fig. 7). Having established the mechanical robustness and operational stability of the hybrid film, we next evaluated its imaging capabilities. To provide a standardized measure of the resolving power of the system, we derived the modulation transfer function (MTF) using the slanted-edge method. The measured ESF (Supplementary Fig. 23a) was differentiated to obtain the line spread function (LSF), which was then fitted with a Gaussian function ($R^2 = 0.92$) (Supplementary Fig. 23b). This analysis yielded an MTF₂₀ (limiting resolution) of 4.1 lp mm^{-1} , surpassing the American College of Radiology (ACR) standard of 2.5 lp mm^{-1} (Supplementary Fig. 23c). This demonstrates that the thin-film geometry effectively prevents lateral light spreading, ensuring sufficient spatial resolution to support structural recovery in the downstream 3D volumetric reconstructions.

To validate the distortion-free imaging performance of our flexible scintillator, we imaged a curved calibration target from multiple viewing angles (Fig. 4b). The target comprised bar patterns of 6 mm, 4 mm, and 2 mm pitch, as well as $13 \text{ mm} \times 13 \text{ mm}$ crossbar arrays on a cylindrical surface. Individual projections were acquired at various orientations (Supplementary Fig. 24). Distortion mapping-based registration confirmed that the multi-angle dataset reproduces the true spatial dimensions of the test patterns, in contrast to the geometric errors observed in single-view acquisition (Fig. 4c).

Multi-angle radiography also provides an effective means to disentangle overlapping structures that are otherwise obscured in planar X-ray projections. Whereas a single frontal view collapses all components into one intensity map, masking critical internal features, acquiring images from rotated perspectives allows individual elements to be resolved. Figure 4d illustrates the experimental configuration for imaging a cylindrical structure containing embedded metal rods, achieved by conformally wrapping a scintillator film around the cylinder surface. Sequential X-ray acquisitions spanning 30° to 210° at 60° intervals produce four complementary radiographs, with individual views capturing either four distinct rods or varying degrees of rod overlap (Fig. 4e). By integrating information across these perspectives and analyzing the spatial distribution of attenuated regions, the underlying rod arrangement can be inferred. Comparison of the computationally derived positions (blue) with the ground-truth rod configuration (yellow) confirmed high positional accuracy (Fig. 4f), with RMS values of 0.798, 0.158, 3.877, and 0.963 for rods 1 through 4, respectively. This enhanced resolution of spatially complex arrangements is particularly critical in medical imaging, for densely packed anatomical structures, demanding precise delineation for reliable diagnosis. For instance, the carpal bones, eight small, tightly clustered bones in the wrist, are often mischaracterized in planar radiography, which has an estimated diagnostic accuracy of approximately 70% compared to CT scans⁶¹. This limitation increases the risk of clinical complications such as non-union, osteonecrosis, and post-traumatic arthritis⁶².

To evaluate the clinical utility of the patch-type scintillator for precise measurement, we performed multi-angle radiography of a chicken tibiotarsus (Fig. 4g). Radiographs acquired at different viewing angles produced distinct morphological variations (Fig. 4h). Furthermore, to evaluate image quality under realistic acquisition conditions, we quantified the contrast-to-noise ratio (CNR) from 32-bit radiographs. Using the mean intensity of the cortical bone ROI ($\mu_{\text{signal}} = 163.9$) and an adjacent background region ($\mu_{\text{background}} = 1.5$, $\sigma_{\text{background}} = 8.3$), we obtained a CNR of 19.5, well above the Rose Criterion (CNR > 5)⁶³ threshold of 5, confirming that the system provides sufficient contrast for reliable structural discrimination under practical imaging conditions. From three sparse-view projections, we extracted sufficient angular and structural information to perform volumetric reconstruction after

masking irrelevant regions and integrating the multi-angle data (Fig. 4i). Also, the projection angles are neither fixed nor restricted to a pre-defined set. The reconstruction operates using the per-view pose and acquisition geometry provided for each capture (Supplementary Fig. 25), allowing full flexibility in angular configuration. Supplementary Fig. 26 quantifies reconstruction quality as a function of angular sparsity by plotting reconstruction metrics versus the number of projection views. The results show that reconstruction accuracy improves as additional angles are incorporated, while a three-view configuration is still sufficient to recover the essential 3D morphology required for portable imaging applications. Beyond angular flexibility, reliable 3D volumetric reconstruction in portable settings requires signal fidelity at reduced radiation doses. To validate this, we further conducted a task-based evaluation of fracture detectability under controlled dose-dependent conditions (Supplementary Fig. 27). As the incident dose decreases, signal contrast diminishes, and the signal-to-noise margin narrows. Our results show that fracture structural discontinuity remains quantitatively distinguishable at baseline exposure (3 mAs, 80 kVp) and at dose levels reduced to ~80% and ~64% of baseline. However, at ~51% of baseline, the CNR drops to 1.9, below the Rose Criterion threshold of 5, and the fracture boundary becomes statistically unresolvable. This establishes a well-defined detectability threshold and confirms that the system supports diagnostically relevant imaging within a dose range compatible with sparse-view acquisition. Such capability is particularly valuable for anatomically complex regions, such as the acetabulum, where accurate classification, preoperative planning, and postoperative monitoring rely on comprehensive structural information that single-view radiography often fails to provide^{63–65}.

Neural network-based 3D reconstruction

Figure 5a shows the workflow for 3D reconstruction, which integrates deep learning-based segmentation with a triplane-based Neural Radiance Field (NeRF) framework using sparse-view X-ray projections (step-by-step visualization in Supplementary Fig. 28). A high-resolution digital hand phantom was placed at the center of a virtual X-ray gantry (Fig. 5b). Three flat-panel detectors at 0° , $+30^\circ$, and -30° each captured a single 2D projection (Views 1–3), emulating clinical constraints on angle count. Diagnostically relevant anatomy was segmented by an ROI network specialized on hand X-ray data. Specifically, we employed the hand gesture identification (MMpose)⁶⁶ algorithm as a Prompt Predictor Network (PPN) with memory-attention blocks to emit positive foreground and negative background point prompts to a frozen SAM image encoder (See methods). Prompts were passed to the SAM mask decoder to produce pixel-accurate masks restricted to clinically meaningful regions¹⁹. The segmented views were then encoded and mapped into a triplane representation for subsequent NeRF-based volumetric reconstruction.

Figure 5c summarizes the network that generates a 3D object. SAM employs a Vision Transformer (ViT)⁶⁷ backbone that tokenizes each X-ray into non-overlapping patches and uses self-attention to relate distant spatial tokens. With point prompts as queries, SAM maintains segmentation consistency across views and produces precise delineate masks that surpass coarse region-based approximations (Fig. 5d and Supplementary Fig. 29), providing a consistent and anatomically precise representation of target structures for the subsequent 3D reconstruction.

The segmented images are then fed into TripoSR²⁰, an implicit model specifically designed to recover high-fidelity 3D volumes from sparse-view without dense volumetric scans (e.g., CT or magnetic resonance imaging (MRI)). Rather than relying on explicit voxel grids or point cloud priors, TripoSR employs a triplane NeRF⁶⁸ representation, decomposing the 3D structure into three orthogonal feature planes (XY, XZ, YZ), to enable efficient volumetric inference from sparse inputs. A structured comparison of representative ultra-sparse reconstruction approaches and their acquisition requirements is

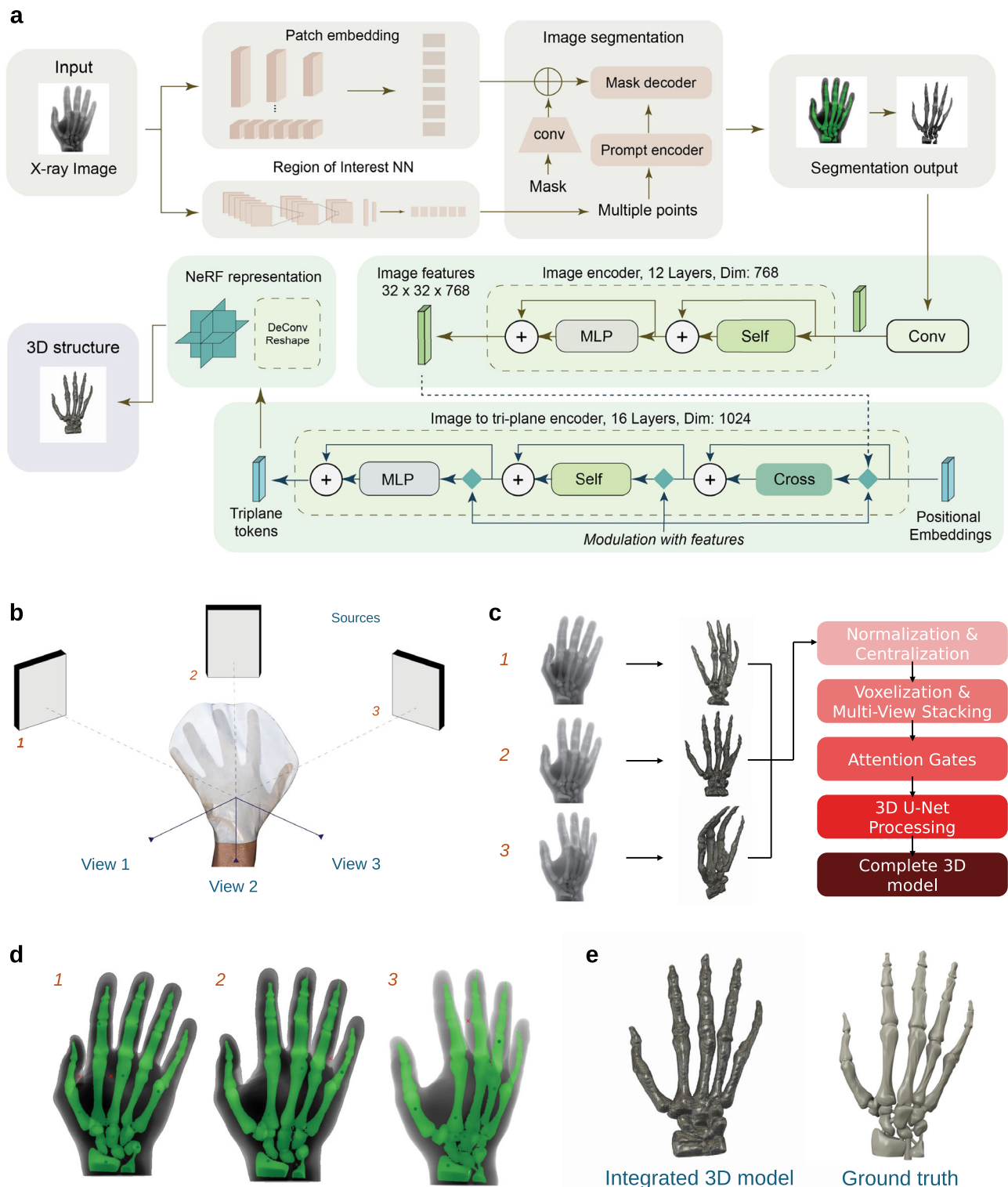


Fig. 5 | Neural network-based segmentation and 3D reconstruction of multi-angle X-ray projections. **a** The overall workflow of the segmentation and 3D reconstruction process. A point-based prompt first delineates the ROI, after which the Segment Anything Model (SAM), comprising patch embedding, prompt encoding, and a mask decoder, produces pixel-level accurate segmentation masks. The resulting segmented images are then processed by TripoSR. A transformer-based image encoder with self-attention and multi-layer perceptron (MLP) modules extracts salient features that are mapped onto a triplane representation via an image-to-triplane encoder. The final 3D structure is reconstructed through a series of deconvolution and reshaping operations. **b** Simulated setup for sparse-view X-ray imaging of a human hand model, in which three projections (View 1-3) were acquired from distinct angles

and used as inputs for neural reconstruction. **c** The corresponding simulated X-ray images (top to bottom: Views 1-3) were segmented using the SAM framework to isolate osseous structures. Individual 3D reconstructions from each view, generated via TripoSR, captured partial geometry that was subsequently normalized to a canonical coordinate frame, voxelized into binary occupancy grids, and fused through a 3D U-Net with attention-gated skip connections, resulting in a complete 3D model. **d** Bone segmentation output obtained using the SAM model with MMPose-based prompt generation, producing a delineated bone map suitable for subsequent processing in the post-3D-generation neural network. **e** In the final fusion step, a volumetric 3D U-Net network integrates intermediate reconstructions through voxel-based attention gating, achieving structural coherence across views.

provided in Supplementary Table T2. Furthermore, by leveraging TripoSR, we utilize the inherent capability of NeRF to reconstruct continuous 3D volumes from sparse, irregularly sampled angles. This computational resilience naturally complements the physical constraints of sparse-view X-ray imaging, transforming a limited set of conformal projections into a dense volumetric model. Radiance F is modeled by $F(x, d) = (\sigma(x), L(x, d))$, where x denotes a 3D point in space, d represents the viewing direction, $\sigma(x)$ is the volumetric density, and $L(x, d)$ represents grayscale intensity at point x . A transformer decoder with cross-attention maps 2D ViT features onto the three triplane, thereby learning consistent spatial correlations across views and enforcing geometric coherence in the reconstructed volume.

A Multi-Layer Perceptron (MLP) decodes triplane features to predict volumetric density and grayscale intensity at any 3D query point. These continuous-valued fields are then composited along each X-ray projection ray via the volumetric rendering integral $I(r) = \int_0^{t_f} T(t)\sigma(r(t))L_s(r(t), d)dt$, where $I(r)$ represents the synthesized intensity, $T(t)$ is the accumulated transmittance modeling X-ray attenuation, σ is the spatial density, and L_s is the radiance emitted at each point along the ray. This continuous representation captures subtle internal density variations (e.g., bones) without explicit surface modeling and preserves fine structural details, including curvature and edge sharpness. Training uses a composite loss function that combines volumetric rendering loss with silhouette consistency loss, and smoothness regularization (see methods), to balance fidelity and geometric regularity.

To reconstruct a complete 3D bone volume from multi-angle radiographs, we fuse view-specific reconstructions through a neural network pipeline. First, each view-specific mesh is normalized into a common coordinate frame centering at the centroid and scaling to fit within $[-0.5, 0.5]^3$. The normalized meshes are voxelized into binary occupancy grids and are then merged by a volumetric 3D U-Net with attention gates, trained to integrate overlapping geometries, fill view-dependent voids and reconcile inconsistencies. A loss function comprised of the Dice coefficient and binary cross-entropy (BCE) was used to ensure that anatomical features and bone morphology are retained in the final output.

For synthetic validation, a 3D anatomical model generated virtual radiographs at three orientations, center frontal (Fig. 5d-2), 30° to the left (Fig. 5d-1) and 30° to the right (Fig. 5d-3). For the network training, we employed more than 5000 datasets generated from multiple anatomically distinct hand models (Supplementary Figs. 30–31). Each projection was first segmented via the SAM model to delineate osseous structures, followed by the reconstruction with TripoSR into three independent 3D density volumes. Finally, the view-specific reconstructions were registered into a common frame and fused through a volume integration network, providing a cohesive 3D representation of the bone anatomy (Fig. 5e). Chamfer Distance (CD) was employed to quantitatively assess the geometric consistency, yielding a value of 0.017 for bone reconstructions. This was supported by a structural similarity index (SSIM) of 0.852 ± 0.037 which measures perceptual similarity in local contrast and structure, the peak signal-to-noise ratio (PSNR) of 19.09 ± 4.52 dB which reflects the relative level of reconstruction error, and the root-mean-square error (RMSE) of 0.127 which captures the average voxel-wise discrepancy between the two volumes, confirming the fidelity of the recovered 3D morphology despite sparse input data^{69,70}. Supplementary Fig. 32 shows the reconstruction accuracy across varying noise levels (Gaussian noise intensity factor 0–10 %) and beam energies (50–80 keV), Supplementary Table 1 and Supplementary Fig. 33 show the system-matched degradation stress test, evaluating reconstruction fidelity under controlled degradation modes, including beam hardening, PSF blur, detector noise, and scatter, applied to simulated projections. Supplementary Fig. 34 further evaluates workflow robustness to out-of-plane (z-axis) angular misalignment across offsets of $\pm 15^\circ$.

In summary, we developed a flexible patch-type perovskite scintillator by hybridizing Cs_3MnBr_5 with a cellulose matrix and seeding nucleation with CNs to promote uniform film growth across a large area. This flexible platform conforms intimately to complex geometries, enabling the acquisition of distortion-free, multi-angle 2D X-ray images that serve as input for AI-driven 3D reconstruction. After identifying the ROIs, the SAM accurately isolates target structures, while the TripoSR network reconstructs these segmented images into high-fidelity 3D models. While modern CT systems provide unparalleled image quality for comprehensive clinical diagnostics, their bulk and infrastructural requirements limit their use in emergency or point-of-care settings. Our integrated platform presents a complementary and versatile alternative. By combining a flexible, conformable scintillator with AI-driven sparse-view reconstruction, it enables high-fidelity 3D imaging in resource-constrained environments where traditional CT is inaccessible. Ultimately, by overcoming the geometric limitations of conventional 2D X-ray imaging, while significantly reducing cumulative radiation exposure compared to CT scans, our integrated AI algorithms with the flexible scintillation platform represent a feasibility-level demonstration of a versatile and patient-friendly approach toward 3D diagnostic imaging applications.

Methods

Fabrication of scintillator patch

The flexible scintillator film is fabricated through a sequential two-step process. First, CNs are uniformly distributed onto a cellulose matrix (No. 6, ADVANTEC) using an electrospray system. A 1 wt% of CNs dispersion in DI water is delivered via syringe pump (EP100, NanoNC) through a single-capillary spinneret with a 22-gauge needle under 21 kV (HV30, NanoNC) at 30 cm tip-to-substrate distance, then dried at 80 °C for over 4 h. Subsequently, Cs_3MnBr_5 microcrystals are grown using a 0.075 mM precursor solution by dissolving 3 mmol of CsBr (99.999%, Sigma-Aldrich) and 1 mmol of MnBr_2 (98%, Sigma-Aldrich) in DI water. The CNs-coated matrix is then immersed in this solution and dried in a convection oven, initially at 50 °C for 6 h, followed by 70 °C for 12 h. Layer-by-layer stacking of CNs-coated cellulose matrix enabled precise control of film thickness. Finally, the hybrid composite is passivated with a PDMS overcoat, Sylgard 184 (DOW) mixed at a 10:1 base-to-curing agent ratio, degassed, applied evenly and cured at 50 °C for 24 h, to yield the encapsulated flexible scintillator film.

Molecular dynamics simulation

Molecular dynamics (MD) simulation was conducted by using Materials Studio 2024. The simulation cell comprised 450 CsBr, 150 MnBr_2 , and 40 cellulose ($\text{C}_6\text{H}_{10}\text{O}_5$) molecules arranged to approximate the early-stage nucleation environment of Cs_3MnBr_5 . Simulations employed the COMPASS III force field, which provides ab initio-derived parameters for both organic and inorganic species. Temperature was maintained at 450 K, corresponding to the crystallization temperature of Cs_3MnBr_5 , using the Nosé–Hoover–Langevin thermostat in the canonical (NVT) ensemble to ensure proper sampling of thermal fluctuations. The integration timestep was set to 1 fs, and each trajectory was propagated for 2 ns. Post-processing analyses were conducted with the Forcite module to extract small-angle X-ray scattering (SAXS) profiles, radial distribution function (RDF), and velocity autocorrelation function (VACF), providing insight into structural ordering and dynamical behavior during hybrid crystallization.

Density functional theory calculations

Density-functional theory (DFT) calculations were performed to evaluate the electronic band structure and interfacial charge transfer. Calculations used the QM/MM simulation cluster [NFEC-2025-02-303332] and Vienna ab initio simulation package (VASP 5.3.5)^{1–3} with the Projector Augmented Wave (PAW) method⁴. The Perdew Burke Ernzerhof (PBE)⁵ form of the generalized gradient approximations

(GGA)⁶ described exchange and correlation. A kinetic energy cutoff of 550 eV was applied for the plane-wave basis set, and electron occupations were treated with Gaussian smearing with a width of 0.05 eV. Structural optimizations were conducted using the conjugate gradient (CG) algorithm⁷ until the residual forces fell below 0.01 eV Å⁻¹.

The tetragonal Cs₃MnBr₅ structure (space group I4/mcm, Materials Project ID: mp-27332)⁸ provided the bulk starting geometry. Relaxed lattice parameters ($a=b=9.70$ Å, $c=15.45$ Å) agreed with experiment to within ~0.7%. A cluster-like thin slab was used to represent the nucleation process of Cs₃

MnBr₅. The (001) facet was chosen due to its low Miller index, as the dominant facet is generally stable and thus tends to be a less active site in crystallization. $3 \times 3 \times 1$ Monkhorst-Pack k-points grids were used for the calculation of slab models for nucleation. The density of states (DOS) calculation showed good agreement with previous first-principles studies⁹. A denser k-mesh of $21 \times 21 \times 13$ was employed for DOS calculations. The formation energies (E_f) are calculated by $E_f = E_{\text{total}} - \sum n \cdot \mu$, where E_{total} is the total energy of Cs₃MnBr₅ with and without cellulose during the nucleation process, μ is the chemical potential of Cs₃MnBr₅ and cellulose, and n is the number of components in the stoichiometry.

To model surface terminations, both Cs–Br-terminated and Br-terminated slabs were constructed, as shown in Supplementary Fig. 35a. A vacuum layer of 20 Å was included to eliminate spurious interactions between periodic images. Increasing the vacuum thickness beyond 20 Å does not affect the calculated work function; therefore, a 20 Å vacuum was used consistently throughout the study. The DOS, as shown in Supplementary Fig. 35b, indicated that the Br-terminated slab better preserved the bulk-like electronic structure. Besides, total energy calculations of two different slabs (Br-terminated: -3.72 eV atom⁻¹ vs Cs–Br-terminated: -3.47 eV atom⁻¹) showed that the Br-terminated slab is energetically more favorable. The thickness of the perovskite slab and the corresponding k-point mesh were chosen based on convergence studies of the DOS (Supplementary Fig. 35c) and work function. Ultimately, a 5-layer symmetric perovskite slab and a Γ -centered k-mesh of $5 \times 5 \times 1$ were used for structural relaxation and self-consistent field (SCF) calculations, while a denser $21 \times 21 \times 1$ mesh was employed for DOS computations.

To determine the optimal interface distance between cellulose and perovskite, the binding energy was calculated as a function of interlayer separation using the Eq. (1) specified by,

$$E_b = \frac{E_{\text{hetero}} - E_{\text{perovskite}} - E_{\text{cellulose}}}{n} \text{ (eV/atom)} \quad (1)$$

Where, n is the total number of atoms, E_{hetero} is the total energy of the heterostructure, $E_{\text{perovskite}}$ is the total energy of isolated Perovskite, and $E_{\text{cellulose}}$ is the total energy of isolated cellulose. As shown in Supplementary Fig. 36, the energy minimum corresponds to an interfacial spacing of approximately 3 Å, indicating the most stable configuration. At this optimal distance, the cellulose layer was further relaxed while keeping the perovskite slab fixed. To incorporate non-local dispersion effects at the interface, the vdW-DF¹⁰ functional was employed, and the cellulose layer was re-relaxed accordingly. Although the inclusion of van der Waals interactions reduced the overall charge transfer, the direction of charge flow remained unchanged.

The charge transfer between perovskite and cellulose was obtained by calculating the electron charge densities of each system before (i.e., $\rho_{\text{cellulose}}$, and $\rho_{\text{perovskite}}$) and after (ρ_{hetero}) contact. The charge density difference, $\Delta\rho$, as a function of position was followed obtained followed by Eq. (2):

$$\Delta\rho = \rho_{\text{hetero}}(Z) - \rho_{\text{cellulose}}(Z) - \rho_{\text{perovskite}}(Z) \quad (2)$$

A planar average was performed to express charge density in 1D, represented by z , the spatial coordinate perpendicular to the interface. The net charge transfer per area (δn) is then evaluated by,

$$\delta n = \int_{z_1}^{z_2} \Delta\rho(z) dz \quad (3)$$

Determining the appropriate limits z_1 and z_2 is critical for accurately capturing the net interfacial charge transfer. A practical criterion was to identify a midpoint z_0 such that:

$$\int_{z_1}^{z_0} \Delta\rho(z) dz = \int_{z_0}^{z_2} \Delta\rho(z) dz \quad (4)$$

In our case, this point was found to be at 28.90 Å. The integrated charge up to this point was calculated as $\delta n = 0.019 e$ per Å², and with a cross-sectional area of 9.70×9.70 Å², the total charge transferred from cellulose to perovskite was approximately 1.7 e per cellulose molecule. When van der Waals interactions were included, the charge transfers were reduced to 1.0 e , while maintaining the same direction of flow. Overall, these results collectively indicate that charge transfer occurs from the cellulose layer to the perovskite.

Characterizations of scintillator patch

For Fourier-transform infrared (FT-IR) spectroscopy, 150 mg powders of both CNs and the hybrid composite were prepared in pellet form. A background spectrum was first acquired, after 64 measurements were averaged to yield the final spectrum. UV-visible absorption measurements were performed at room temperature using an HP 8453 UV-vis spectrometer (Agilent) with a 10 mm quartz. Prior to analysis, Cs₃MnBr₅ dissolved in DI water was dispersed in isopropyl alcohol (IPA), and the solvent baseline was subtracted. Spectras were acquired from 200 to 800 nm range at a scan speed of 10 nm s⁻¹. For PL measurements, a 405 nm diode laser (L405-5W, Deloslaser) was coupled to a microscope through a single-mode optical fiber (PI-305A-FC, Thorlabs) and positioned via a piezoelectric stage (Albatross, Nanofocus). The excitation beam was split 50:50 (BSW10R, Thorlabs) and filtered by a long-pass edge filter (FF02-409/LP-25, Semrock). Time-correlated single-photon counting (TCSPC) measurements used a 375 nm picosecond diode laser (BDS-SM Lasers, Becker & Hickl) via a free space mirror mount system. Laser intensities were modulated with a variable neutral density (ND) filter, and emission was collected through a multi-mode optical fiber (M50L02S-A, Thorlabs) into a spectrometer (SR-3031-B, Andor) with PL spectra acquired via SOLIS software (Andor), and lifetime data was recorded using custom-built software. The microscope setup and white light illumination was controlled by NANOimage software (Nanofocus Inc.). Mechanical robustness was assessed by subjecting the films, mounting on an elastic support, to 10 mm bending at a speed of 2 mm s⁻¹ using an STS-RT-1AXIS (Sciencetown). After a predefined number of bending cycles, PL from the bent region was measured to assess performance retention. Radioluminescence (RL) spectra were conducted under ambient conditions using an X-RAD 320 system (Precision Inc.) operated at 200 kVp and 10 mA and acquired with a spectrometer (CCS200, Thorlabs) through a 200 μ m core optical fiber.

Radioluminescence measurements

Radioluminescence spectra were acquired across the diagnostic X-ray range (50–250 kVp, 5–40 mA) using a X-ray irradiator (X-RAD 320TM, Precision, USA). The flexible scintillator film was placed orthogonally to the incident beam, and the emitted light was collected via a 200 μ m core optical fiber (Thorlabs) into a compact fiber-coupled spectrometer (CCS200, Thorlabs), covering 200–1000 nm wavelength range. All measurements were performed at room temperature under ambient conditions. To verify the delivered dose, Gafchromic™ EBT3

radiochromic films (Ashland Inc., USA) were placed immediately behind the scintillator to quantify absorbed radiation through optical density analysis of the exposed films.

X-ray Imaging procedure

X-ray radiography was performed using a custom imaging module under diagnostic settings of 80 kVp and 3 mA with a 1 second exposure (EXSYS Portable, DRTEC). Image acquisition was carried out with a flat-panel TFT detector (model EXT 2430H, DRTECH), with the flexible scintillator film placed directly on top of the detector surface. Data acquisition and processing were conducted using the ECalil software. For three-dimensional reconstruction, projections were acquired at multiple orientations (0° , $\pm 30^\circ$). This sparse-view strategy substantially reduces cumulative radiation exposure by reconstructing volumetric data from only three projections. Imaging performance was assessed by measuring spatial resolution with opaque patterns, evaluating edge sharpness, and performing structural visualization across the various viewing angles.

Neural network algorithm

Physics-based synthetic data generation. To overcome the scarcity of paired volumetric-radiographic clinical data, we established a high-fidelity simulation pipeline (Supplementary Fig. 1) utilizing the gVirtualXR (gVXR). This approach enabled the deterministic generation of large-scale training datasets with ground-truth geometric precision. To ensure anatomical consistency across the synthetic samples, we employed various distinct 3D hand models (Supplementary Fig. 37), which were normalized to a fixed sagittal height of 200 mm and centrally aligned within the simulation volume. We employed a parallel-beam configuration to emulate far-field acquisition, thereby minimizing perspective distortion and simplifying subsequent geometric calibration. The virtual detector was modeled as a planar sensor array with 1440×1440 pixels and a pixel pitch of 0.2 mm, establishing a fixed spatial sampling frequency. Radiographic synthesis utilized mono-chromatic beams with energies spanning 0.05–0.08 MeV, and photon statistics were modeled by constraining the flux to 500–1000 photons per ray, simulating varying dose regimes.

Each simulated scene comprised a dual-component phantom consisting of registered surface meshes for skeletal and soft-tissue structures. The bone component was assigned to a multi-element cortical-bone composition at a density of 1.85 g cm^{-3} , while soft tissue was approximated as a water-equivalent homogeneous medium with a density of 1.0 g cm^{-3} . Photon transport was computed via ray-casting, where intensity reduction followed the Beer–Lambert law of exponential attenuation based on the cumulative path length through these distinct material densities. To foster generalization in the reconstruction pipeline, we generated a heterogeneous dataset by varying ray energies, photon noise statistics, and anatomical morphologies. For each hand model, multi-angle projections were acquired by stochastically sampling azimuthal angles across a full angular range. This process yielded a paired dataset connecting 2D radiographic projections with their corresponding 3D volumetric ground-truth, providing a robust foundation for supervised learning.

ROI identification Network. The ROI identification pipeline was designed to isolate diagnostically relevant regions in X-ray images and generate point-based prompts for subsequent segmentation using the SAM. First, raw X-ray images were enhanced using gamma correction to improve bone visibility. Images were normalized to $[0, 1]$, transformed via logarithmic scaling for better dynamic range, and then gamma-corrected ($\gamma=0.5$) to brighten mid-tones while preserving bone contrast. The resulting images were inverted so that bone regions appear as high-intensity areas suitable for downstream processing. Next, MMPose, a pre-trained hand pose estimation network, was employed to detect 21 anatomical keypoints corresponding to the

hand skeletal structure (wrist, finger joints, and fingertips). The model outputs keypoint coordinates (x, y) along with confidence scores for each detected landmark. These coordinates were used for the prompt generation for SAM: we converted detected keypoints into positive and negative point prompts for SAM-based bone segmentation:

Positive prompts include: (1) all detected keypoints except fingertips (16 points), and (2) interpolated points between the wrist and each finger base to ensure coverage of metacarpal bone regions. Interpolation factors ($\alpha=0.45\text{--}0.55$) were empirically tuned to place prompts along bone shafts.

Negative prompts include: (1) extrapolated points beyond each fingertip (extending 10% past the detected tip along the finger direction) to exclude non-bone regions, and (2) external boundary points placed at a fixed margin (50 pixels) outside the hand bounding box (top, left, right, and corner positions) to explicitly mark background regions.

Segment anything model. The Segment Anything Model (SAM) was applied to perform bone segmentation using point-based prompts generated by the ROI network. Image encoder of SAM was built on a Vision Transformer (ViT-Huge) backbone and comprised 32 transformer blocks, each with a hidden dimension of 1280 and 16 attention heads. Input X-ray images are tokenized into 16×16 spatial patches, enabling the self-attention layers to capture both local and global contextual features. Point prompts, annotating positive (bone) and negative (background) pixels, are converted by the prompts encoder into positional embeddings and incorporated as auxiliary tokens alongside the image tokens. A lightweight transformer-based mask decoder, consisting of 2 layers and 256 hidden dimensions each, then produces a 256×256 pixel-wise segmentation mask using a sigmoid activation. We used the model that has been trained with a composite loss function combining focal loss and Dice loss, which were employed to mitigate class imbalance and improve segmentation accuracy. The focal loss is defined as,

$$\mathcal{L}_{\text{Focal}} = -\frac{1}{N} \sum_{i=1}^N \alpha_i (1 - p_i)^\gamma \log(p_i) \quad (5)$$

where α_i is the class weighting factor (0.75 for bone, 0.25 for background). $\gamma (=2.0)$ is the focusing parameter, p_i is the predicted probability for pixel i , and N is the total number of pixels. The Dice loss is given by

$$\mathcal{L}_{\text{Dice}} = 1 - \frac{2 \sum y_i p_i}{\sum y_i + \sum p_i} \quad (6)$$

where y_i is the ground truth label. The overall segmentation loss is defined as Eq. (7)

$$\mathcal{L}_{\text{Seg}} = \mathcal{L}_{\text{Focal}} + \lambda \mathcal{L}_{\text{Dice}} \quad (7)$$

with $\lambda = 1.0$. The SAM used the AdamW optimizer with a learning rate of 5×10^{-5} , weight decay of 1×10^{-4} , and a batch size of 4. A linear decay schedule was applied to the learning rate, and the network was fine-tuned for 50 epochs.

TripoSr 3D Reconstruction Network. The TripoSr model employs a ViT encoder consisting of 12 transformer blocks with a hidden dimension of 768, and 12 attention heads. Segmented X-ray images are first resized to 512×512 pixels and tokenized into non-overlapping 16×16 patches. These patches were subsequently processed by the transformer encoder to extract 2D feature representations. These encoded features are fed into a transformer-based decoder incorporating cross-attention layers which project them onto three orthogonal triplane feature grids, each with a resolution of 64×64 and a feature

dimension of 40. The decoder comprises 16 transformer layers, each with 16 attention heads and a hidden dimension of 1024, alternating self-attention over the triplane tokens with cross-attention to the encoder features.

With X-ray image input, the model simulates X-ray formation by tracing rays through a volumetric domain and sampling radiometric and structural information at discrete points along each ray. At each sample location, the 3D location onto three orthogonal 2D feature maps (triplanes), which capture learned structural details from different angles. These combined features are then passed to a multilayer perceptron (MLP) head, comprising 10 hidden layers of 64 units and SiLU activation, which predicts volumetric density and grayscale intensity at each queried 3D coordinate. Final 2D projections are synthesized via volumetric rendering for the prediction of X-ray attenuation (grayscale). The model was trained by a composite loss function, incorporating MSE rendering loss, learned perceptual image patch similarity (LPIPS) perceptual loss, and binary cross entropy (BCE) loss. The MSE rendering loss is the pixel-wise L2 difference between the synthetic projection and the corresponding real X-ray image:

$$\mathcal{L}_{\text{MSE}} = \frac{1}{N} \sum_i \|I_{\text{pred},i} - I_{\text{gt},i}\|^2 \quad (8)$$

where $I_{\text{pred},i}$ and $I_{\text{gt},i}$ are the predicted and ground truth grayscale intensities at pixel i , and N is the number of pixels. The LPIPS perceptual loss compares the rendered projection and the ground truth image in the deep feature space of a pre-trained VGG network rather than at the pixel level, providing a learned perceptual similarity measure that better aligns with human judgments of image quality than pixel-wise losses:

$$\mathcal{L}_{\text{LPIPS}} = \sum_l \frac{1}{H_l W_l} \sum_{h,w} \|w_l \odot (\hat{\phi}_l(I_{\text{pred}})_{h,w}) - \hat{\phi}_l(I_{\text{gt}})_{h,w}\|_2^2 \quad (9)$$

where $\phi_l(\cdot)$ extracts features at layer l , H and W are the spatial dimensions of feature map. w_l is the learned channel-wise weight vector at l , and $\odot$ stands for element-wise (Hadamard) multiplication. The BCE compares the rendered foreground silhouette to the ground-truth silhouette:

$$\mathcal{L}_{\text{BCE}} = -\frac{1}{N} \sum_i [M_{\text{gt},i} \log(M_{\text{pred},i}) + (1 - M_{\text{gt},i}) \log(1 - M_{\text{pred},i})] \quad (10)$$

where M_i is the rendered mask probability at pixel i , M_{gt} is the ground-truth mask, and N is the number of pixels. The final loss function was the weighted sum²⁰:

$$\mathcal{L}_{\text{TripoSR}} = \mathcal{L}_{\text{MSE}} + \lambda_{\text{LPIPS}} \mathcal{L}_{\text{LPIPS}} + \lambda_{\text{BCE}} \mathcal{L}_{\text{BCE}} \quad (11)$$

with $\lambda_{\text{LPIPS}} = 2.0$ and $\lambda_{\text{BCE}} = 0.05$.

3D U-Net fusion network

Multi-view 3D reconstructions were fused using a 3D U-Net with attention-gated skip connections. Each view-specific mesh from TripoSR was normalized to a canonical coordinate frame $[-0.5, 0.5]^3$ and voxelized into a binary occupancy grid at resolution $192 \times 192 \times 192$. The N-view voxel volumes were stacked as input channels to the network. The encoder path consisted of four stages with progressively increasing feature channels (16-32-64-128-256), each employing two $3 \times 3 \times 3$ convolutions with batch normalization and ReLU activation, followed by max pooling. The decoder path used transposed convolutions for upsampling, with attention gates applied to skip

connections to selectively emphasize geometrically consistent features across views. The output layer produced voxel-wise occupancy probabilities via sigmoid activation. The network was trained using a combined Dice and BCE loss:

$$\mathcal{L} = \lambda_{\text{Dice}} \mathcal{L}_{\text{Dice}} + \lambda_{\text{BCE}} \mathcal{L}_{\text{BCE}} \quad (12)$$

where:

$$\mathcal{L}_{\text{Dice}} = 1 - \frac{2 \sum_i p_i g_i + \epsilon}{\sum_i p_i + \sum_i g_i + \epsilon} \quad (13)$$

$$\mathcal{L}_{\text{BCE}} = -\frac{1}{N} \sum_i [g_i \log(p_i) + (1 - g_i) \log(1 - p_i)] \quad (14)$$

Here, p_i denotes the predicted occupancy probability and g_i the ground truth label for voxel i . The weighting factors were set to $\lambda_{\text{Dice}} = 0.5$ and $\lambda_{\text{BCE}} = 0.5$. A small constant $\epsilon = 1 \times 10^{-6}$ was added for numerical stability. Optimization was performed using AdamW optimization (learning rate: 1×10^{-3} , weight decay: 1×10^{-4}).

Evaluation protocols and training/test parameters

To ensure complete reproducibility, reconstruction performance is evaluated on a held-out test split generated via a fixed random seed. The dataset, organized by sample identifiers, contains multiple radiographic projections and their corresponding ground-truth 3D bone geometries. We employ a deterministic 80/10/10 split (train/validation/test), ensuring all reported metrics represent the full test distribution rather than a curated subset. For each test case, the sparse input is constructed by sampling three views from predefined angular acquisition windows (-180° to -135° , -45° to 45° , and 135° to 180°), with a minimum angular separation constraint enforced to avoid near-duplicate views.

Prior to 3D volumetric estimation, all 2D projection images are preprocessed. Intensities are normalized, log-scaled to compress dynamic range, gamma-corrected ($\gamma = 0.5$), and inverted to enhance cortical bone contrast. To isolate the target anatomy and suppress background structure, 21 anatomical key-points are estimated and converted into positive and negative point prompts for the SAM. The resulting segmentation mask gates each radiograph to a consistent region of interest (ROI). This preprocessing and ROI-gating strategy is applied identically across training and testing phases to prevent data leakage.

To avoid artificially inflating performance metrics, reconstruction outputs and ground-truth volumes are evaluated within a deterministic canonical coordinate frame defined by the acquisition geometry. Both prediction and target are converted into discrete voxel occupancy grids utilizing identical coordinate normalization and alignment parameters. Per-case iterative closest point (ICP) registration is deliberately omitted prior to metric computation, as volumes share a common canonical frame, ICP is unnecessary, and its application would risk concealing intrinsic geometric failure modes and rotational instabilities. Quantitative evaluation is performed in the 3D voxel domain using a fixed grid resolution and static binarization thresholds, eliminating post-processing-dependent variance. When surface-based metrics are required, voxel occupancies are converted into point clouds using a fixed occupancy threshold (0.5) and a capped sampling budget (up to 10,000 points), with coordinates normalized to $[-1, 1]$. Standard mesh extraction, smoothing, and decimation are utilized exclusively for qualitative visualization. All reported quantitative metrics are derived directly from the fixed voxel and point-cloud conversion rules described above.

Volumetric fidelity is quantified using Peak Signal-to-Noise Ratio (PSNR) and Root Mean Square Error (RMSE), computed directly in

the 3D voxel domain. PSNR uses a fixed dynamic range with a maximum value of 1.0. Structural Similarity Index (SSIM) is evaluated in 3D using a localized Gaussian window (size 7, $\sigma = 1.5$) with standard stability constants. Macroscopic geometric accuracy is quantified via the Chamfer distance between thresholded voxel-derived point clouds^{11,12}, calculated as the average bidirectional nearest-neighbor distances subject to the sampling budget and spatial normalization described above.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

The synthetic X-ray and mesh dataset generated in this study has been deposited in the Zenodo repository under DOI 10.5281/zenodo.19798835. The dataset includes one hand bone 3D mesh model, which is publicly available. The remaining four hand bone mesh models are not publicly available as they are derived from commercially available samples that require consent for distribution. Access to these additional meshes can be obtained by contacting the corresponding author; they will be distributed under restricted access after completion of a data use agreement stipulating that the data will not be redistributed and will be used solely for reproducing the results of this study. Access to these additional meshes can be obtained by contacting the corresponding author. Requests will be responded to within 14 business days. Upon approval, a data use agreement must be completed stipulating that the data will not be redistributed and will be used solely for reproducing and extending the results of this study. Once access is granted, the data will remain available for the duration of the research project. These restrictions apply to all requestors equally, regardless of institutional affiliation.

Code availability

The source code for the full pipeline, including X-ray simulation, segmentation, per-view 3D reconstruction, multi-view fusion training, and evaluation, is publicly available at https://github.com/bb8ax/xray_to_voxl/ under the MIT License.

References

- Wang, Y. et al. Efficient X-ray luminescence imaging with ultra-stable and eco-friendly copper(I)-iodide cluster microcubes. *Light Sci. Appl.* **12**, 155 (2023).
- Zhu, W. et al. Low-dose real-time X-ray imaging with nontoxic double perovskite scintillators. *Light Sci. Appl.* **9**, 112 (2020).
- Gan, N. et al. Organic phosphorescent scintillation from copolymers by X-ray irradiation. *Nat. Commun.* **13**, 3995 (2022).
- Guzelturk, B. et al. Bright and durable scintillation from colloidal quantum shells. *Nat. Commun.* **15**, 4274 (2024).
- Chen, H. et al. Clinical characteristics and intrauterine vertical transmission potential of COVID-19 infection in nine pregnant women: a retrospective review of medical records. *Lancet* **395**, 809–815 (2020).
- Li, H. et al. Damaged lung gas exchange function of discharged COVID-19 patients detected by hyperpolarized Xe MRI. *Sci. Adv.* **7**, eabc8180 (2021).
- Reilly, R. B. et al. An altered natural killer cell immunophenotype characterizes clinically severe pediatric RSV infection. *Sci. Transl. Med.* **16**, eado6606 (2024).
- Consiglieri, G. et al. Early skeletal outcomes after hematopoietic stem and progenitor cell gene therapy for Hurler syndrome. *Sci. Transl. Med.* **16**, eadi8214 (2024).
- Hu, X. et al. Density-discriminating chromatic x-ray imaging based on metal halide nanocrystal scintillators. *Sci. Adv.* **9**, eadh5081 (2023).
- Battle, C. et al. Comparison of the use of lung ultrasound and chest radiography in the diagnosis of rib fractures: A systematic review. *Emerg. Med. J.* **36**, 185–190 (2019).
- Hansen, A. E. et al. Multimodal soft tissue markers for bridging high-resolution diagnostic imaging with therapeutic intervention. *Sci. Adv.* **6**, eabb5353 (2020).
- Zhang, Y. et al. Unsupervised X-ray image segmentation with task driven generative adversarial networks. *Med. Image Anal.* **62**, 101664 (2020).
- Ma, J. et al. Segment anything in medical images. *Nat. Commun.* **15**, 654 (2024).
- Fok, W. Y. R. et al. Adversarial robustness improvement for X-ray bone segmentation using synthetic data created from computed tomography scans. *Sci. Rep.* **14**, 25813 (2024).
- Shiode, R. et al. 2D–3D reconstruction of distal forearm bone from actual X-ray images of the wrist using convolutional neural networks. *Sci. Rep.* **11**, 15249 (2021).
- Kim, D. et al. Accurate auto-labeling of chest X-ray images based on quantitative similarity to an explainable AI model. *Nat. Commun.* **13**, 1867 (2022).
- Saporta, A. et al. Benchmarking saliency methods for chest X-ray interpretation. *Nat. Mach. Intell.* **4**, 867–878 (2022).
- Fan, W. et al. A deep-learning-based framework for identifying and localizing multiple abnormalities and assessing cardiomegaly in chest X-ray. *Nat. Commun.* **15**, 1347 (2024).
- Kirillov, A. et al. “Segment Anything”. In *Proceedings of the IEEE/CVF international conference on computer vision* (2023).
- Tochilkin, D. et al. TripoSR: Fast 3D object reconstruction from a single image. Preprint at <https://doi.org/10.48550/ARXIV.2403.02151> (2024).
- Pan, W. et al. Cs₂AgBiBr₆ single-crystal X-ray detectors with a low detection limit. *Nat. Photonics* **11**, 726–732 (2017).
- Kim, Y. C. et al. Printable organometallic perovskite enables large-area, low-dose X-ray imaging. *Nature* **550**, 87–91 (2017).
- Ding, X. et al. Preparation of CsPbBr₃Films for Efficient Perovskite Solar Cells from Aqueous Solutions. *ACS Appl. Energy Mater.* **4**, 5504–5510 (2021).
- Dutta, S. et al. Simple water-assisted synthesis of Ce-doped lead-free cesium metal halides for visible-blind UV photodetectors. *ACS Sustain. Chem. Eng.* **11**, 92–100 (2023).
- Dutta, S. et al. Pure tunable emissions from cesium manganese bromide by monitoring the crystal fields through a sustainable approach. *Adv. Sustain. Syst.* **8**, 2400092 (2024).
- Vighnesh, K. et al. Hot-injection synthesis protocol for green-emitting cesium lead bromide perovskite nanocrystals. *ACS Nano* **16**, 19618–19625 (2022).
- Yuan, Z. et al. Hot-electron injection in a sandwiched TiO_x-Au-TiO_x structure for high-performance planar perovskite solar cells. *Adv. Energy Mater.* **5**, 1500038 (2015).
- Baek, K. Y. et al. Mechanochemistry-driven engineering of 0D/3D heterostructure for designing highly luminescent Cs–Pb–Br perovskites. *Nat. Commun.* **13**, 4263 (2022).
- Nanbakhsh, K. et al. On the longevity and inherent hermeticity of silicon-ICs: evaluation of bare-die and PDMS-coated ICs after accelerated aging and implantation studies. *Nat. Commun.* **16**, 12 (2025).
- Choi, E. Y. et al. Development of moisture-proof polydimethylsiloxane/aluminum oxide film and stability improvement of perovskite solar cells using the film. *RSC Adv.* **9**, 11737–11744 (2019).
- Ingham, B. et al. How nanoparticles coalesce: An in situ study of Au nanoparticle aggregation and grain growth. *Chem. Mater.* **23**, 3312–3317 (2011).
- Demmel, F. et al. Diffusion in liquid metals is directed by competing collective modes. *Phys. Rev. B.* **111**, L081104 (2025).

33. Shiinoki, M. et al. Molecular dynamics simulation of diffusion behavior in liquid Sn and Pb. *Metall. Mater. Trans. B Process Metall. Mater. Process. Sci.* **55**, 278–286 (2024).
34. Atkyan, N. et al. Raman and FT-IR spectroscopy investigation the cellulose structural differences from bacteria *Gluconacetobacter sucrofermentans* during the different regimes of cultivation on a molasses media. *AMB Express* **10**, 84 (2020).
35. Wulandari, W. T. et al. Nanocellulose prepared by acid hydrolysis of isolated cellulose from sugarcane bagasse. *IOP Conf. Ser. Mater. Sci. Eng.* **107**, 012045 (2016).
36. Nghiem, L. D. et al. Role of electrostatic interactions in the retention of pharmaceutically active contaminants by a loose nanofiltration membrane. *J. Membr. Sci.* **286**, 52–59 (2006).
37. Zhang, C. et al. Surface modification on cellulose nanofibers by TiO₂ coating for achieving high capture efficiency of nanoparticles. *Coatings* **9**, 139 (2019).
38. Cho, J. et al. Membrane filtration of natural organic matter: factors and mechanisms affecting rejection and flux decline with charged ultrafiltration (UF) membrane. *J. Membr. Sci.* **164**, 89–110 (2000).
39. Xu, M. et al. Heating revival of Cs₃MnBr₅ for anti-counterfeiting and large-area flexible X-ray imaging. *Opt. Mater.* **156**, 115959 (2024).
40. Li, K. et al. Regulating the optical properties of Cs₃MnBr₅ nanocrystals in glasses for narrow-band green emission. *Inorg. Chem.* **62**, 13001–13010 (2023).
41. Chen, Q. et al. All-inorganic perovskite nanocrystal scintillators. *Nature* **561**, 88–93 (2018).
42. Rodnyi, P. A. *Physical Processes in Inorganic Scintillators*. (CRC press, 2020).
43. Wang, Y. et al. Perovskite scintillators for improved X-ray detection and imaging. *Angew. Chem. Int. Ed.* **62**, e202304638 (2023).
44. Li, F. Halide perovskites, a game changer for future medical imaging technology. *Biophys. Rev.* **6**, 011302 (2025).
45. Gupta, S. K. et al. Nanoscintillators. in *21st Century Nanoscience—A Handbook* 1–6 (CRC Press, 2020).
46. Ju, Y. et al. Lead-free manganese halide perovskite for minute-level dynamic time-gating anticounterfeiting. *ACS Appl. Mater. Interfaces* **16**, 70774–70782 (2024).
47. Kwon, S. B. et al. Organic solvent-free lyophilization assisted recrystallization synthesis of high-purity green emissive Cs₃MnX₅ (X = I, Br). *J. Alloys Compd.* **845**, 156324 (2020).
48. Peng, H. et al. Evolution of the structure and properties of mechanochemically synthesized pyrrolidine incorporated manganese bromide powders. *J. Mater. Chem. C.* **8**, 6488–6495 (2020).
49. Ma, W. et al. Thermally activated delayed fluorescence (TADF) organic molecules for efficient X-ray scintillation and imaging. *Nat. Mater.* **21**, 210–216 (2022).
50. Wolszczak, W. W. et al. “Toward perovskite-related scintillators with necessary Stokes shift and thickness for hard X-ray radiography and gamma spectroscopy”. in *Advanced X-ray Detector Technologies: Design and Applications* (Springer, Cham, 2022).
51. Alqorashi, A. K. et al. DFT insights into the stoichiometric (001), (011) and (111) thin film surfaces of Ba₂NalO₆. *Solid State Commun.* **387**, 115537 (2024).
52. Haruyama, J., Sodeyama, K., Han, L. & Tateyama, Y. Termination dependence of tetragonal CH₃ NH₃ PbI₃ surfaces for perovskite solar cells. *J. Phys. Chem. Lett.* **5**, 2903–2909 (2014).
53. Haruyama, J., Sodeyama, K., Han, L. & Tateyama, Y. Surface properties of CH₃ NH₃ PbI₃ for perovskite solar cells. *Acc. Chem. Res.* **49**, 554–561 (2016).
54. Schulz, P., Cahen, D. & Kahn, A. Halide perovskites: is it all about the interfaces? *Chem. Rev.* **119**, 3349–3417 (2019).
55. Eglitis, R. I. & Jia, R. Review of systematic tendencies in (001), (011) and (111) surfaces using B3PW as well as B3LYP computations of BaTiO₃, CaTiO₃, PbTiO₃, SrTiO₃, BaZrO₃, CaZrO₃, PbZrO₃ and SrZrO₃ Perovskites. *Materials* **16**, 7623 (2023).
56. Kong, Q. et al. Phase engineering of cesium manganese bromides nanocrystals with color-tunable emission. *Angew. Chem. Int. Ed.* **60**, 19653–19659 (2021).
57. Nagarkar, V. V. et al. Structured CsI(Tl) scintillators for X-ray imaging applications. In *Proceedings 1997 IEEE Nuclear Science Symposium Conference Record*. (1997).
58. Derenzo, S. E. et al. Experimental efforts and results in finding new heavy scintillators. In *Proceedings of the Crystal 2000 International Workshop on Heavy Scintillators for Scientific and Industrial Applications*. (1992).
59. Popelar, T. et al. Universal radiative lifetimes in the long-lived luminescence of Si quantum dots. *J. Phys. Chem. C.* **127**, 20426–20437 (2023).
60. Senden, T. et al. Photonic effects on the radiative decay rate and luminescence quantum yield of doped nanocrystals. *ACS Nano* **9**, 1801–1808 (2015).
61. You, J. S. et al. The usefulness of CT for patients with carpal bone fractures in the emergency department. *Emerg. Med. J.* **24**, 248–250 (2007).
62. Gray, R. R. L. et al. Scaphoid fracture and nonunion: new directions. *J. Hand. Surg. Eur. Vol.* **48**, 4S–10S (2023).
63. Kohli, A. et al. Three-dimensional CT and 3D MRI of Hip- important aids to Hip preservation surgery. *Semin. Ultrasound CT MRI* **44**, 252–270 (2023).
64. Scheinfeld, M. H. et al. Acetabular fractures: What radiologists should know and how 3D CT can Aid Classification. *Radiographics* **35**, 555–577 (2015).
65. Verbeek, D. O. et al. Postoperative CT is superior for acetabular fracture reduction assessment and reliably predicts hip survivorship. *J. Bone Jt. Surg. Am.* **99**, 1745–1752 (2017).
66. MMPose Contributors. OpenMMLab Pose Estimation Toolbox and Benchmark. (2020).
67. Dosovitskiy, A. et al. An image is worth 16x16 words: transformers for image recognition at scale. Preprint at <https://doi.org/10.48550/arXiv.2010.11929> (2020).
68. Mildenhall, B. et al. NeRF: Representing scenes as neural radiance fields for view synthesis. *Commun. ACM* **65**, 99–106 (2022).
69. Quach, M., Chetouani, A., Valenzise, G. & Dufaux, F. A deep perceptual metric for 3D point clouds. *Electron. Imaging* **33**, 257–1–257–7 (2021).
70. Nilsson, J. & Akenine-Möller, T. Understanding SSIM. Preprint at <https://doi.org/10.48550/arXiv.2006.13846> (2020).

Acknowledgements

SSD and MZ acknowledge the UVA Rivanna HPC Cluster. This work is dedicated to the memory of Hyunsik Im.

Author contributions

Conceptualization: D.H.P., J.C., Seokho K. Methodology: K.L., K.K., S.C., Seokho K, B.B., and J.C. Investigation: Sungwoo K, J.K., J.B., C.B.L., and Y.J.L. Visualization: S.C., Siyeon K, B.B., Y.B., and Y.J.L. Computational analysis: C.B.L., S.S.D., M.Z., K.K., and B.B. Funding acquisition: K.L., D.H.P., and Seokho K. Project administration: K.L., D.H.P., and Seokho K. Supervision: D.H.P. and K.L. Writing – original draft: K.L., Seokho K, B.B., Y.B., D.W.K., and Y.J.L. Writing – review & editing: K.L., K.K., S.C., M.Z., and Y.J.L.

Funding

U.S. Air Force Office of Scientific Research (AFOSR) Young Investigator Program (YIP) (grant no. FA9550-23-1-0159, K.L.). National Science Foundation (NSF) (grant no. ECCS-2332060, K.L.). National Research Foundation of Korea (NRF) grant, funded by the Korea government Ministry of Science and ICT (MSIT) (RS-2024-00357783, D.H.P.). Korea Institute for Advancement of Technology (KIAT) grant funded by the Korea Government (MOTIE) (P0020460, DHP). National Science

Foundation (NSF) (grant no. 2421213, M.Z.). National Research Foundation of Korea (NRF) grant, funded by the Korea government Ministry of Science and ICT (MSIT) (RS-2024-00463084, S.K., Y.B.). Engineering Research Center (ERC), funded by the National Research Foundation of Korea (RS-2022-NR070840, K.K.). This research has been supported by the POSCO Science Fellowship of POSCO TJ Park Foundation (S.C.).

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41467-026-73320-9>.

Correspondence and requests for materials should be addressed to Kyeounghak Kim, Sangeun Cho, Dong Hyuk Park or Kyusang Lee.

Peer review information *Nature Communications* thanks Xiaogang Liu, Xiaowang Liu and Zhehui Wang for their contribution to the peer review of this work. A peer review file is available.

Reprints and permissions information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2026