OEA | Opto-Electronic Advances

Article

May 2026, Vol. 9, No. 5

Open Access

AI-powered nonlinear optical imaging reveals protein spatial homogenization as an indicator of impaired bone quality in type 2 diabetes

Bowen Zhang[1†], Jiangbo Pu[1†], Tao Hu[2], Junjie Zeng[3], Han Zhang[3], Zemeng Chen[1], Xiang Ji[1], Shuhua Yue[3], Lin Z. Li[4] and Ting Li[1*]

Abstract: Type 2 diabetes mellitus (T2DM) significantly elevates fracture risk, a severe complication often underestimated by conventional bone mineral density (BMD) assessments. Here, we applied label-free multimodal nonlinear optical (NLO) imaging with AI-powered texture feature analysis to characterize T2DM-related bone quality alterations. Our results identified aberrant spatial protein distribution, characterized by increased homogeneity and reduced contrast, as a distinctive pathological feature in T2DM bone. The alterations in spatial distribution were also observed in hydroxyapatite (HA) and autofluorescent metabolites. A K-nearest neighbor (KNN) model, trained on fused texture features from these three components, achieved a superior classification accuracy of 93.56% in distinguishing T2DM-related bone tissues, markedly outperforming single-component models (~70%). This demonstrated that fused multi-component spatial distribution features offer enhanced discriminative power for quantifying T2DM-associated pathological changes. Collectively, aberrant molecular spatial distribution, particularly of protein, represents a potentially unappreciated indicator of diabetic bone quality alterations. Integrating multimodal NLO imaging with explainable AI offers a novel approach for unraveling the mechanistic underpinnings of complex pathological alterations, which not only overcomes the limitations of conventional biomarker assessment but also establishes a powerful framework for discovering new pathological targets.

Keywords: label-free nonlinear optical image; type 2 diabetes mellitus; impaired bone quality; explainable AI; multimodal integration

DOI: 10.29026/oea.2026.250312 | CSTR: 32247.14.oea.2026.250312

Citation: Zhang BW, Pu JB, Hu T et al. AI-powered nonlinear optical imaging reveals protein spatial homogenization as an indicator of impaired bone quality in type 2 diabetes. *Opto-Electron Adv* **9**, 250312 (2026).

1 Introduction

Type 2 diabetes mellitus (T2DM), the predominant form of diabetes and a major global health crisis, affected nearly all of the approximately 529 million individuals living with diabetes worldwide in 2021[1]. Beyond well-established complications such as nephropathy and retinopathy[1–3], the increased fracture risk has also emerged as a prevalent and severe complication of T2DM, posing significant clinical challenges[4–7]. While bone mineral density (BMD) remains the current gold standard for bone quality assessment[8], it frequently underestimates fracture risk in T2DM patients. Emerging evidence highlights that alterations in skeletal microstructure and material composition are key pathological drivers of T2DM-related fractures[9,10]. However, these precise pathomorphological characteristics underlying these changes remain elusive.

Current techniques for imaging these alterations rely

Received: 11 November 2025

Accepted: 9 March 2026

Published online: 19 March 2026

[1]Biomedical Engineering Institute, Chinese Academy of Medical Sciences and Peking Union Medical College, Tianjin 300192, China; [2]Department of Spine Surgery, Shanghai East Hospital, School of Medicine, Tongji University, Shanghai 200092, China; [3]Key Laboratory of Biomechanics and Mechanobiology (Beihang University), Ministry of Education, Institute of Medical Photonics, Beijing Advanced Innovation Center for Biomedical Engineering, School of Biological Science and Medical Engineering, Beihang University, Beijing 100191, China; [4]Britton Chance Laboratory of Redox Imaging, Department of Radiology, Perelman School of Medicine University of Pennsylvania, Philadelphia, PA 19104, USA.

[†]These authors contributed equally to this work.

*Correspondence: T Li, E-mail: liting@bme.cams.cn

heavily on histological staining, including light microscopy[11,12] and confocal microscopy[13,14]. These methods necessitate time-consuming decalcification and complex staining protocols, often leading to low success rates and the inability to comprehensively label multiple key constituents simultaneously. Such limitations collectively hinder their application in identifying T2DM-related bone pathology. Therefore, there is a critical need for a high-resolution imaging modality capable of rapid, label-free, and comprehensive quantitative analysis of bone microstructure and composition to elucidate the underlying mechanisms of diabetic bone fragility.

Multimodal label-free nonlinear optical (NLO) microscopy offers a powerful solution to these challenges, enabling simultaneous visualization of multiple constituents without the need for decalcification or staining[15]. By leveraging the intrinsic vibrational spectroscopy of various molecular components, NLO imaging (e.g., stimulated Raman scattering (SRS)[16], second harmonic generation (SHG)[17], and two-photon excited fluorescence (TPEF)[18]) allows for quantitative, high-resolution imaging of intact bone microstructure and composition in situ[19–21]. SRS aligns with spontaneous Raman scattering spectroscopically, providing quantitative distribution of molecular concentration[22]. SHG and TPEF are particularly valuable in capturing collagen fibers[23,17,24] and auto-fluorescent metabolites[25], respectively. Thus, the integration of these complementary modalities (SRS/SHG/TPEF) is essential for comprehensively characterizing the complex histological features of bone tissue[26–28], which is uniquely suited for probing the microstructural and compositional underpinnings of T2DM-related bone fragility[29–31,32].

In T2DM patients, vertebrae and hip are the most susceptible sites for fractures[2,33], with vertebral fractures linked to a higher mortality rate. The osteon, as the primary structural unit of cortical bone in vertebrae, plays a critical role in determining bone quality[34]. Consequently, our study focused on vertebral cortical bone, utilizing the osteon as the basic unit of investigation to ensure comparability. In structure, osteons consist of Haversian canals, lamellae, osteocytes, and the lacunocanalicular network (LCN)[35,36]. The Haversian canal, located at the center of the osteon, is encircled by multiple layers of lamellae[37]. Osteocytes, the sole cell type found in mature osteons, reside within the lamellae[38], interconnected by dendrite-like canaliculi forming the LCN[39]. In terms of material composition, osteons are primarily composed of lipids, proteins, hydroxyapatite (HA), and collagen fibers. Haversian canals contain blood vessels, nerves, and lipid-rich bone marrow. The multilayer lamellae consist of interconnected collagen fibers, with calcium and phosphorus deposited to maintain bone strength[40,41]. Osteocytes, rich in signaling lipids and structural proteins, and the LCN are crucial for signal transmission[42] and regulating bone remodeling and turnover, all contributing to skeletal quality[43].

In this study, we utilized label-free multimodal NLO imaging (integrating SRS, SHG, and TPEF) to characterize T2DM-related alterations in osteon microstructure and material composition. We employed texture extraction and Artificial Intelligence (AI) algorithms to identify these changes from five channels, including proteins, lipids, HA, collagen fibers, and auto-fluorescent metabolites. The results revealed that T2DM was associated with distinct alterations in multi-component distribution, especially protein characterized by increased homogeneity and decreased contrast in skeletal texture compared to controls. Although single-component texture features provided limited discriminative power (~70% accuracy), the fusion of multiple channels dramatically improved recognition accuracy to 93.56%. This work highlights the potential significance of multi-component spatial distribution in understanding T2DM-related impaired bone quality.

2 Results

2.1 Multimodal NLO imaging of osteons

Linear decomposition of lipid and protein. Multimodal images of osteons in the 4th or 5th lumbar vertebrae (L4 or L5) were acquired using three NLO techniques, including SRS, SHG, and TPEF. Leveraging the close alignment of SRS spectra with spontaneous Raman spectra, we employed the peak frequencies in the CH_2 (2850 cm^{-1}), CH_3 (2930 cm^{-1}), and PO_4^{-3} (959 cm^{-1}) bands to image lipids, proteins, and HA, respectively[44]. Then, the premeasured calibration matrix was used to perform linear decomposition of SRS images at 2850 cm^{-1} and 2930 cm^{-1} to map lipid and protein distribution (Fig. 1(a, b))[45]. This approach addressed decreased mapping precision from the overlapping Raman spectra between protein and lipid standards at 2850 cm^{-1} and 2930 cm^{-1}. Because the Raman shift of HA is distinct from the CH stretching band, HA channel did not require the linear decomposition. The details of the decomposition formula are provided in the Methods section.

Raw SRS images distinctly delineated the fundamental structures of osteons, including the Haversian canal, osteocytes, bone lamellae, and LCNs (Fig. S2). Unexpectedly, vertebral osteons displayed irregular morphologies, differing from the conventional round or elliptical form in long bones[46,47]. Critically, T2DM patients had decreased quantity and connectivity of bone canaliculi compared to controls (the patients without T2DM), suggesting potentially impaired biomechanical signal transmission between osteocytes (Fig. 1(a, b)). Although raw 2930 cm^{-1} images showed higher signal intensities than 2850 cm^{-1} (lipid) images, post-decomposition images revealed comparable overall protein and lipid concentrations (Fig. 1(d, f)). However, two-color merged images highlighted differences in their spatial distribution (Fig. 1). In control osteons, proteins were predominantly localized in the bone lacunae, canaliculi, and central

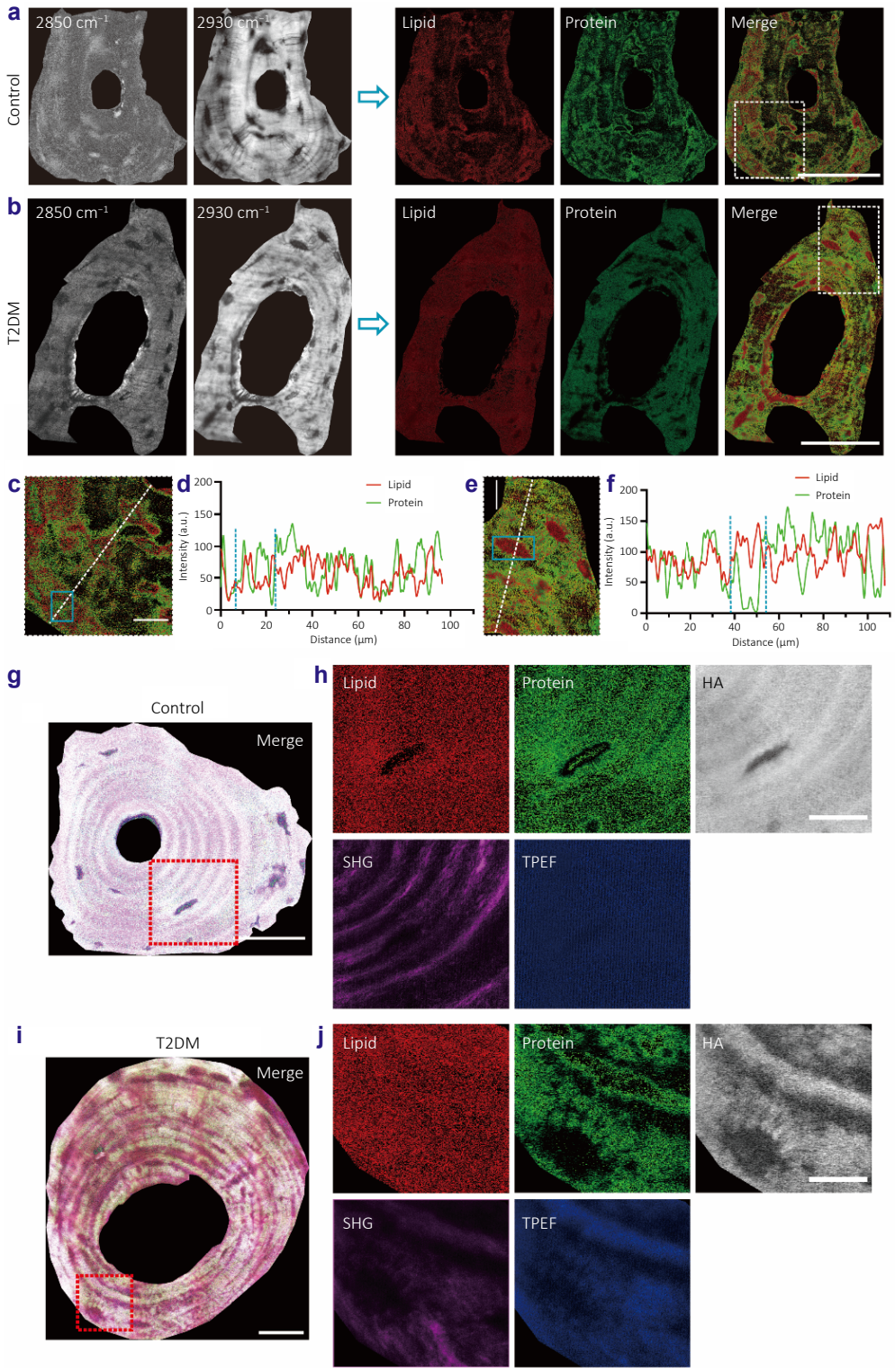

Fig. 1 | NLO imaging of osteons in control and T2DM groups. (a, b) Representative SRS images of one osteon in control (a) and T2DM (b) groups and decomposed images of lipids (red), proteins (green), and overlap (Merge). Scale bar, 100 μm. (c) Enlarged detail image of the white rectangle in (a). Scale bar, 20 μm. (d) A line profile of lipids and proteins in (c) (the white dotted line), and two blue dashed lines delineated the intensity profile within the solid blue box in (b). (e) Enlarged detail image of the white rectangle in (b). Scale bar, 20 μm. (f) A line profile of osteocyte in (e) (the dotted line), and blue dashed lines delineated the intensity profile within the blue box in (e). (g) Representative five-color composited images showing the distribution of lipids (red), proteins (green), HA (gray), SHG (magenta), and TPEF (blue) in one osteon of no T2DM group. Scale bar, 50 μm. (h) Detailed images of five channels in the red rectangle of (g). Scale bar, 20 μm. (i) Representative five-color images showing lipids (red), proteins (green), HA (gray), SHG (magenta), and TPEF (blue) in one osteon of the T2DM group. Scale bar, 50 μm. (j) Detailed images of five channels in the red rectangle of (i). Scale bar, 20 μm.

regions of osteocytes, while lipids primarily surrounded osteocytes (Fig. 1(c, d), blue box). Conversely, T2DM osteons displayed osteocytes largely enveloped by lipids, with proteins predominantly confined to bone lamellae (Fig. 1(e, f)), indicating a distinctive reorganization of osteocytes and lamellar material composition.

Multiple channels quantitative analysis. Following linear decomposition, we merged five channels from SRS, SHG, and TPEF to further investigate the material distribution changes within osteons. The resultant five-color composite images revealed distinct spatial distributions of multiple materials: HA and collagen fibers, along with auto-fluorescent components, predominantly localized in the extracellular matrix (ECM), while lipids accumulated mainly in osteocytes and proteins predominantly in the LCNs and osteocytes (Fig. 1(g–j)). This pattern, aligned with osteon structure, underscored the efficacy of multimodal NLO techniques in imaging the complex architecture of osteons in cortical bone. Furthermore, we conducted a preliminary quantitative analysis to discern differences in osteons between two groups. We observed that protein, lipid, and HA channels showed similar concentrations and intensity distributions across both groups (Fig. S5(a)), whereas T2DM osteons exhibited notably higher skewness and kurtosis in the SHG channel and reduced values in the TPEF channel (Fig. S5(b, c)). This indicated the substantial potential of NLO signal distribution features in identifying T2DM-related bone changes.

Orientation analysis of collagen fiber. SHG imaging provided intricate details about the collagen fibers within circumferential lamellae (Fig. 2(a, b))[24,48]. Given the established link between altered collagen fiber structure and increased bone fragility[17], we hypothesized that changes in collagen composition or orientation might contribute to mechanical alterations observed in T2DM bones. We found no significant difference in the intensity of SHG image signals between the two groups (Fig. S5(a)), revealing that collagen fiber content was not the primary pathological alteration in T2DM cortical bone. To further evaluate collagen fiber organization, we quantified orientation distribution, local energy, and local coherency. The coherency value ranged between 0 and 1, with 0 indicating isotropic areas and 1 indicating highly oriented structures[49]. The pixel block with higher energy value corresponded to higher anisotropic structures. Analysis of SHG images, alongside corresponding energy maps (Fig. 2(c, d)) and coherency (Fig. 2(e, f)) maps revealed that regions of higher coherency and energy predominantly aligned with well-defined textures in both groups. However, no significant difference in collagen fiber orientation was observed between T2DM and control groups (Fig. 2(g–i)). Thus, these results collectively suggest that alterations in collagen fiber content or overall orientation may not be a primary contributor to the impaired bone quality related to T2DM in this context.

2.2 Machine learning for distinguishing T2DM osteons

Although qualitative analysis of NLO images hinted at differences in material distribution within osteons between T2DM and control groups, these observations were not robustly supported by single-channel statistical metrics. This underscored the inadequacy of an individual imaging channel or single feature to fully capture the subtle, but critical pathological nuances. To overcome this limitation, we applied Machine Learning (ML) techniques to distinguish T2DM-affected osteons, leveraging high-dimensional features extracted from multimodal NLO images.

Texture features have been identified as powerful descriptors in image analysis and application[50–52]. In radiomics, first-order statistical features, shape-based features, and gray-level related second-order texture features[50] played pivotal roles for identifying lesions in clinical imaging examinations, such as Computed Tomography (CT)[53], and Magnetic Resonance Imaging (MRI)[54,55]. Given the parallels between radiological images and multimodal NLO images, our study employed them to elucidate T2DM-induced osteon changes.

Our analysis workflow included five pivotal stages: image preprocessing, texture feature extraction, single channel classification, and two distinct channel fusion strategies (feature concatenation and a modified stacking ensemble algorithm), as detailed in Fig. 3. To enhance the robustness and generalizability of our ML models and to optimally capture the integral loop-like structure of osteons, raw images underwent a transformation into a polar coordinate system and were divided into eight sub-images, complemented by data augmentation. We then extracted global first-order statistical features and local spatial gray-level co-occurrence matrix (GLCM) features for classification[56–58]. Shape-based features were excluded, as T2DM-related bone quality decline is not primarily reflected in gross morphological lesions.

Classification with single channel features. To evaluate the classification performance of texture features across individual imaging channels and to identify optimal algorithms, we employed eight conventional ML models for single-channel classification of osteons between the two groups. Across five imaging channels, protein and TPEF channels achieved the two highest classification performances across most algorithms and feature sets (base texture and GLCM features at five different distances, Tables 1–2 and Tables S2–S5). Within the same feature set, K-Nearest Neighbor (KNN) and ensemble algorithms, such as Random Forest (RF), Adaptive Boosting (AdaBoost), and Gradient Boosting Decision Tree (GBDT) generally demonstrated superior adaptability. Notably, the highest average accuracy (77.55%±4.06%) for single-channel classification was derived from KNN model utilizing base texture features from protein channel (sensitivity 65.11%±9.48%, specificity 86.04%±5.63%, area under curve (AUC) 0.82±0.05, Table 1).

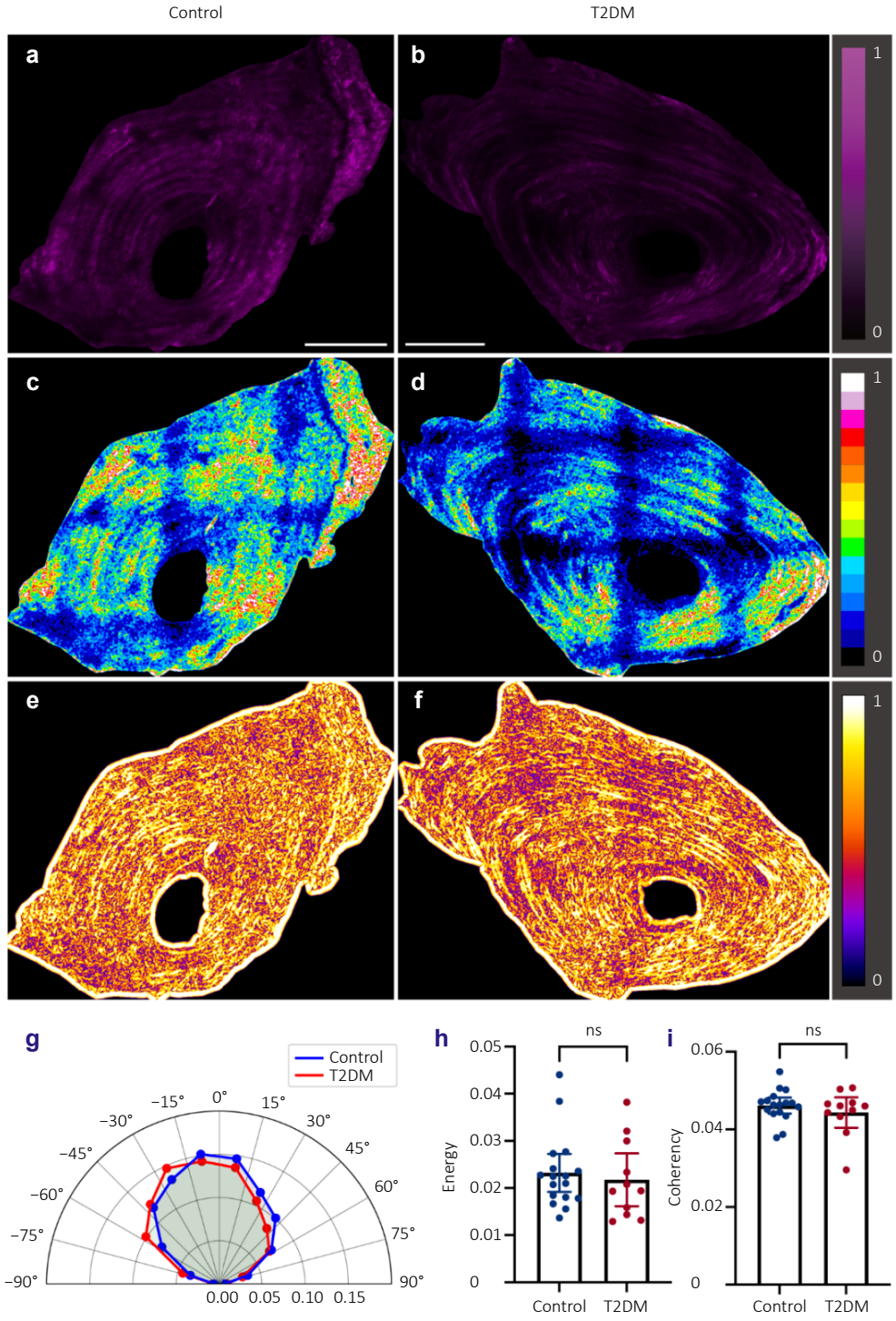

Fig. 2 | The orientation of collagen fibers in osteons. (a, b) Representative SHG images of osteon in control (a) and T2DM (b) groups. (c, d) Local energy maps of SHG images in (a) and (b), respectively. (e, f) Local coherency maps of SHG images in (a) and (b), respectively. The basic calculation element is the 40×40 pixel block. (g) Average local orientation distribution of osteons in control and T2DM groups. The orientation ranged from −90° to 90°, and every 15° was as a statistical interval. (h) Average local energy of one osteon in SHG image. (i) Average local coherency of one osteon in SHG image. Data were shown as mean with 95% CI. n = 17 in control group, n = 11 in T2DM group. ns p>0.05 by two-sided Student's t-test. Scale bar, 50 μm.

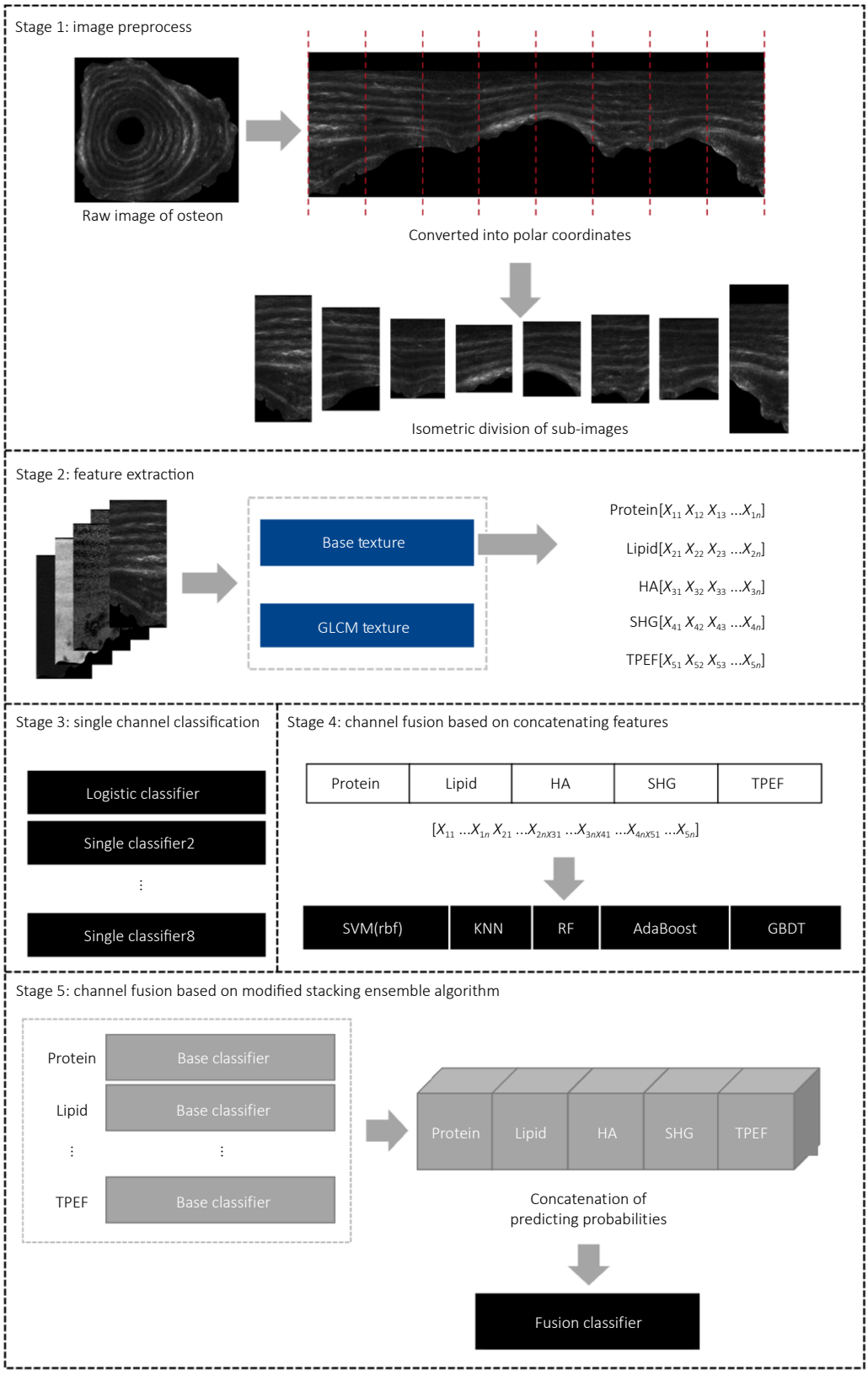

Fig. 3 | Study workflow for distinguishing osteons in T2DM and control groups by machine learning algorithms.

Although base texture features marginally outperformed GLCM features, neither achieved robust classification, underscoring the inherent challenge in comprehensively characterizing T2DM-induced bone microstructure and material composition changes using single-channel features alone.

Channel fusion based on concatenating features. We employed two strategies for channel fusion to enhance classification performance, including feature concatenation and modified stacking ensemble algorithm (Stages 4 and 5 in Fig. 3). For feature concatenation, we concatenated texture features extracted from 5 channels to form the higher-dimensional features as the input of models. Four algorithms with relatively high classification accuracy in single-channel classification (maximum accuracy > 80%) were selected to classify these concatenated features. The widely used SVM model was also included as a baseline reference for accuracy comparisons, despite not meeting the inclusion criteria.

We found that feature concatenation substantially improved the classification performance for both base texture and GLCM compared to single-channel approaches, including recognition accuracy and model stability. Except for the baseline model (SVM(rbf)), the remaining four algorithms achieved accuracies consistently above 80% and AUC values exceeding 0.9 across all tests (Fig. 4(a, b)). Consistent with single-channel results, KNN remained the most effective algorithm for recognizing T2DM-related osteon changes, achieving an AUC value of 0.97±0.02 with base texture features and 0.98±0.02 with GLCM features (D = 4 pixels). The three tree-based ensemble algorithms (RF, AdaBoost, and GBDT) exhibited comparable categorization performance, outperforming the baseline SVM(rbf) but performing lower than KNN (Fig. 4(d)).

Evaluating GLCM features performance across various distances revealed that shorter-distance GLCM features (D = 1, 2, 4 pixels) exhibited superior identification power compared to longer-distance GLCM features (D = 8, 16 pixels), indicating higher sensitivity to T2DM-affected osteon changes (Fig. 4(d), Fig. S7(a)). Integrating other metrics of confusion matrix, including specificity, sensitivity, and AUC (Fig. S7(a) and S7(c)), we identified 4-pixel distance as the optimal GLCM hyperparameter for classification in this dataset and thereby subsequently employed it

Table 1 | Single classification performance of base texture features in different algorithms.

No.	Channel	ACC(%)	SEN(%)	SPE(%)	AUC	No.	Channel	ACC(%)	SEN(%)	SPE(%)	AUC
Logistic						**RF**					
1	HA	61.85±4.75	10.30±5.71	95.80±5.82	0.65±0.06	**21**	**HA**	**68.48±3.63**	**53.03±9.13**	**79.25±7.83**	**0.76±0.05**
2	SHG	62.60±5.43	09.58±4.62	97.09±3.27	0.57±0.06	22	SHG	60.55±5.01	26.18±8.55	83.48±6.57	0.61±0.06
3	TPEF	63.81±6.70	23.28±12.66	91.42±7.12	0.73±0.06	**23**	**TPEF**	**74.81±4.62**	**60.05±8.78**	**84.72±6.82**	**0.79±0.05**
4	Protein	71.53±5.16	34.93±9.26	95.59±3.73	0.70±0.06	**24**	**Protein**	**73.72±4.31**	**50.35±9.51**	**89.36±7.16**	**0.82±0.05**
5	Lipid	59.72±5.67	00.37±1.30	98.37±3.80	0.43±0.05	25	Lipid	67.85±5.36	51.02±8.68	79.35±7.56	0.71±0.07
Bayes						**GBDT**					
6	HA	60.76±5.51	14.19±6.38	91.16±5.94	0.60±0.07	26	HA	68.29±3.75	53.95±8.83	78.14±7.44	0.75±0.04
7	SHG	59.62±6.11	14.97±5.72	88.47±4.71	0.59±0.05	27	SHG	62.44±5.85	36.89±9.96	79.52±6.84	0.62±0.06
8	TPEF	58.42±6.31	79.12±15.83	45.70±15.33	0.71±0.06	**28**	**TPEF**	**74.27±4.14**	**60.95±7.65**	**83.12±6.42**	**0.79±0.05**
9	Protein	71.35±4.98	35.10±8.45	94.90±4.06	0.70±0.05	**29**	**Protein**	**74.90±4.00**	**58.40±8.95**	**86.00±5.88**	**0.82±0.04**
10	Lipid	52.82±7.51	26.56±24.83	70.17±26.45	0.42±0.06	30	Lipid	66.76±5.32	51.04±9.07	77.45±8.08	0.71±0.07
SVM(rbf)						**AdaBoost**					
11	HA	63.28±5.32	12.49±7.23	96.73±4.55	0.73±0.06	31	HA	68.22±3.95	58.15±8.90	75.12±6.61	0.74±0.05
12	SHG	63.30±5.58	07.79±4.51	99.27±1.29	0.59±0.07	32	SHG	61.11±5.55	38.44±9.83	76.22±6.23	0.60±0.06
13	TPEF	65.00±6.53	33.15±18.36	87.57±9.45	0.74±0.06	**33**	**TPEF**	**72.36±4.36**	**62.90±7.89**	**78.79±7.06**	**0.78±0.05**
14	Protein	73.09±5.02	32.02±8.34	99.79±1.18	0.69±0.05	**34**	**Protein**	**73.74±4.00**	**62.67±8.35**	**81.12±5.53**	**0.82±0.04**
15	Lipid	59.77±5.61	00.85±1.68	98.14±3.36	0.54±0.06	35	Lipid	66.03±5.13	53.51±9.00	74.58±8.35	0.70±0.07
KNN						**Decision tree**					
16	**HA**	**73.35±3.91**	**67.46±7.14**	**77.38±5.72**	**0.77±0.05**	36	HA	63.47±4.30	51.63±12.88	71.80±9.63	0.69±0.05
17	SHG	58.01±5.25	39.56±8.74	70.19±6.01	0.57±0.06	37	SHG	58.76±4.96	35.00±10.98	74.88±8.63	0.57±0.06
18	**TPEF**	**71.28±5.19**	**56.34±9.09**	**81.24±7.41**	**0.76±0.05**	38	TPEF	71.22±4.64	58.30±8.09	79.92±7.30	0.72±0.05
19	**Protein**	**77.55±4.06**	**65.11±9.48**	**86.04±5.63**	**0.82±0.05**	**39**	**Protein**	**72.11±4.24**	**50.87±12.21**	**86.43±8.70**	**0.77±0.05**
20	Lipid	68.97±4.91	63.96±9.99	72.33±6.42	0.74±0.07	40	Lipid	64.94±5.11	50.27±9.62	74.85±7.80	0.66±0.06

Logistic Logistic regression, Bayes Naïve Bayesian, SVM(rbf) rbf-kernel support vector machine, KNN K-Nearest Neighbor, RF random forest, AdaBoost Adaptive Boosting, GBDT Gradient Boosting Decision Tree, GLCM Gray-Level Co-occurrence Matrix, ACC Accuracy, SEN Sensitivity, SPE Specificity, AUC, the Area Under the Curve.

Table 2 | Single classification performance of GLCM features (D = 4 pixels) in different algorithms.

No.	Channel	ACC(%)	SEN(%)	SPE(%)	AUC	No.	Channel	ACC(%)	SEN(%)	SPE(%)	AUC
Logistic						**RF**					
1	HA	63.00±5.26	17.52±8.81	93.36±6.95	0.68±0.05	21	HA	65.40±5.07	48.48±9.48	76.83±7.52	0.72±0.06
2	SHG	61.42±5.65	12.92±5.18	93.01±4.09	0.61±0.06	22	SHG	59.46±4.47	25.11±7.95	82.32±6.89	0.56±0.05
3	TPEF	58.50±4.85	20.68±9.85	83.92±7.40	0.65±0.06	**23**	**TPEF**	**73.03±4.45**	**61.98±8.58**	**80.63±7.13**	**0.80±0.05**
4	Protein	72.62±4.32	40.19±8.11	93.93±5.51	0.71±0.05	**24**	**Protein**	**72.00±4.08**	**53.04±10.94**	**84.91±7.10**	**0.78±0.04**
5	Lipid	62.64±5.83	09.09±4.92	97.67±3.78	0.55±0.06	25	Lipid	69.00±5.01	49.52±10.23	82.48±6.84	0.72±0.06
Bayes						**GBDT**					
6	HA	61.85±5.54	14.10±5.38	92.95±4.23	0.64±0.08	26	HA	65.09±5.10	50.84±10.63	74.68±8.17	0.70±0.06
7	SHG	59.71±5.65	16.89±5.53	87.47±4.99	0.55±0.07	27	SHG	57.51±4.02	32.16±7.07	74.59±6.61	0.56±0.06
8	TPEF	57.81±5.93	19.12±8.52	83.19±6.28	0.65±0.06	**28**	**TPEF**	**72.54±4.57**	**61.60±8.67**	**79.86±6.72**	**0.80±0.05**
9	Protein	73.10±4.59	39.29±7.82	94.97±4.75	0.66±0.06	**29**	**Protein**	**71.76±4.33**	**54.85±10.33**	**83.24±7.32**	**0.77±0.05**
10	Lipid	62.09±5.57	22.70±6.86	87.99±8.42	0.57±0.07	30	Lipid	68.39±5.07	51.37±10.02	80.09±6.69	0.72±0.06
SVM(rbf)						**AdaBoost**					
11	HA	61.73±4.76	19.25±13.63	90.56±11.53	0.70±0.06	31	HA	61.01±4.62	48.06±10.22	69.90±6.90	0.65±0.06
12	SHG	63.61±5.40	11.16±5.08	97.71±2.95	0.59±0.07	32	SHG	58.84±5.04	38.75±7.94	72.31±7.76	0.59±0.06
13	TPEF	63.53±5.34	43.33±20.54	78.92±13.56	0.74±0.06	**33**	**TPEF**	**71.70±4.88**	**61.31±8.75**	**78.52±7.36**	**0.79±0.05**
14	**Protein**	72.81±4.72	35.53±7.81	96.93±3.27	0.75±0.05	**34**	**Protein**	69.17±4.10	55.63±9.35	78.43±7.19	0.74±0.05
15	Lipid	62.39±5.50	09.42±4.47	96.87±4.25	0.61±0.09	35	Lipid	70.20±4.30	57.43±7.93	78.99±6.92	0.72±0.05
KNN						**Decision tree**					
16	**HA**	65.74±4.66	59.28±10.45	69.97±5.93	0.70±0.07	36	HA	63.41±4.13	53.96±11.17	70.33±7.79	0.66±0.05
17	SHG	58.21±3.87	48.16±6.85	65.29±5.95	0.58±0.05	37	SHG	57.51±4.15	30.58±8.84	75.29±7.00	0.55±0.05
18	**TPEF**	**70.77±4.30**	**65.07±7.87**	**74.64±6.20**	**0.76±0.05**	**38**	TPEF	69.28±4.70	57.42±12.06	77.32±7.12	0.74±0.05
19	**Protein**	**71.78±5.17**	**62.47±9.25**	**78.03±6.82**	**0.75±0.06**	**39**	**Protein**	**69.45±4.36**	**51.08±9.28**	**81.94±7.74**	**0.71±0.05**
20	Lipid	64.73±4.74	53.52±8.64	72.23±6.59	0.67±0.05	40	Lipid	62.84±4.51	44.76±10.41	75.24±8.52	0.65±0.06

Logistic Logistic regression, Bayes Naïve Bayesian, SVM(rbf) rbf-kernel support vector machine, KNN K-Nearest Neighbor, RF random forest, AdaBoost Adaptive Boosting, GBDT Gradient Boosting Decision Tree, GLCM Gray-Level Co-occurrence Matrix, ACC Accuracy, SEN Sensitivity, SPE Specificity, AUC, the Area Under the Curve.

in feature concatenation analysis.

Furthermore, compared to base texture derived from first-order statistics, the second-order GLCM features with spatial information significantly enhanced the classification performance (Fig. 4(e), Fig. S7(b)). For KNN model, GLCM features (D = 4 pixels) consistently achieved accuracies predominantly above 90%, whereas base texture accuracies generally ranged from 80% to 90%. However, combining base texture and GLCM features did not result in additional improvements in classification performance for most algorithms (Fig. 4(b) vs. Fig. S7(d)), indicating that GLCM features had already encapsulated the information present in base texture. The concatenation of both feature types only led to information redundancy without enhancing discriminatory power. Consequently, for Stage 4, we identified GLCM (D = 4 pixels) as the most appropriate feature set and the KNN classifier as the most effective algorithm for discriminating T2DM-induced osteon changes.

To determine the specific differences in multi-channel texture features distribution between the two groups, we applied Z-score normalization and visualized features in matrices (Fig. 4(c)). We found that T2DM osteons predominantly exhibited texture features distributed towards the extremes (darker in red and blue), while control osteons displayed distributions centered around the mean (lighter red and blue). Both base texture and GLCM features of protein channel showed the most remarkable differences between the two groups compared to other channels. Although high-dimensional channel feature concatenation provided a dramatic increase in accuracy, it was equally crucial to identify explainable T2DM-associated bone structural and material alterations. Such insights not only advance our understanding of pathological mechanisms but also serve as a foundational basis for the clinical diagnosis of T2DM-related fractures.

Feature importance analysis. To ascertain major influential channels and features within fusion features, we employed three complementary methods. For three tree-based ensemble algorithms, we derived feature importance (Gini importance) based on model structure, summarizing the contribution of each channel (Fig. 5(a) base texture, Fig. 5(b) GLCM (D = 4 pixels)). In GBDT and RF models, protein channel features consistently dominated, contributing over one-third of the total importance, surpassing TPEF

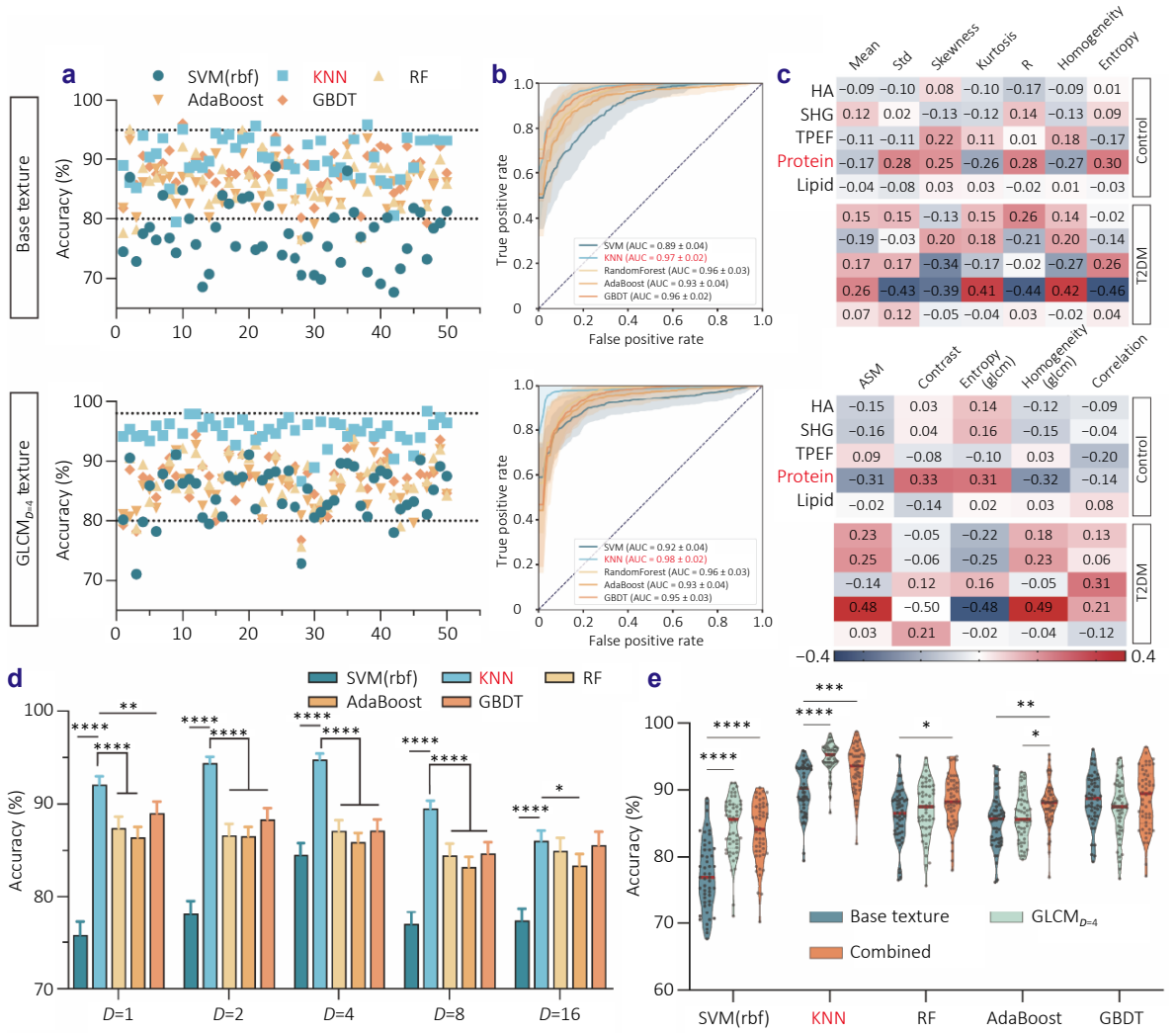

Fig. 4 | Classification performance of texture feature fusion. (a, b) Classification accuracies (from 50 random train-test splits, a) and corresponding ROC curves of five machine learning models for fused base texture (upper) and GLCM (D = 4 pixels, lower) features. (c) Min-max normalized base texture (upper) and GLCM (D = 4 pixels, lower) feature matrix of multiple channels in both control and T2DM groups. (d) The classification accuracy of fused GLCM features from average direction of five distances (D = 1, 2, 4, 8, 16 pixels). Bars represent the average accuracy with 95% CI (from 50 random train-test splits). (e) Multiple comparisons of performance for five models across fused base texture, GLCM (D = 4 pixels) and their combined features. The red line marked the mean accuracy. Data were shown in (D) as mean with 95% CI. *p<0.05, **p<0.01, and ****p<0.0001 by Two-way ANOVA with Tukey's *post-hoc* test.

(~1/5) and lipid (~1/5) channels (Fig. S8(a, c)). Conversely, the remaining four channels (excluding SHG) showed more uniform importance in AdaBoost model, irrespective of feature type (Fig. S8(b)). HA-related texture features exhibited variable classification importance dependent across models, which was associated with specific model structures. In contrast, SHG channel consistently demonstrated the lowest contribution to classification across all models.

To further dissect the specific distributions, we also ranked the top 10 most important individual features in the prediction process of each model. For base texture, the mean, entropy, homogeneity of protein channel, alongside the mean of lipid channel, emerged as consensus features of three models, playing pivotal roles in classification (Fig. 5(a_1–a_3)). For GLCM (D = 4 pixels) features, protein homogeneity and contrast, lipid homogeneity, and TPEF correlation and entropy were identified as particularly important (Fig. 5(b_1–b_3)). However, a broader comparison of feature importance across different algorithms revealed that the classification performance of tree-based models was not solely contingent on any single feature (Fig. S9). This suggested that the spatial variations observed in different components were heterogeneous and could not be adequately captured by a single characteristic alone.

Furthermore, to comprehensively assess the individual and synergistic contributions of each channel on classification performance, we implemented a stepwise channel selection approach. This method not only simplified model

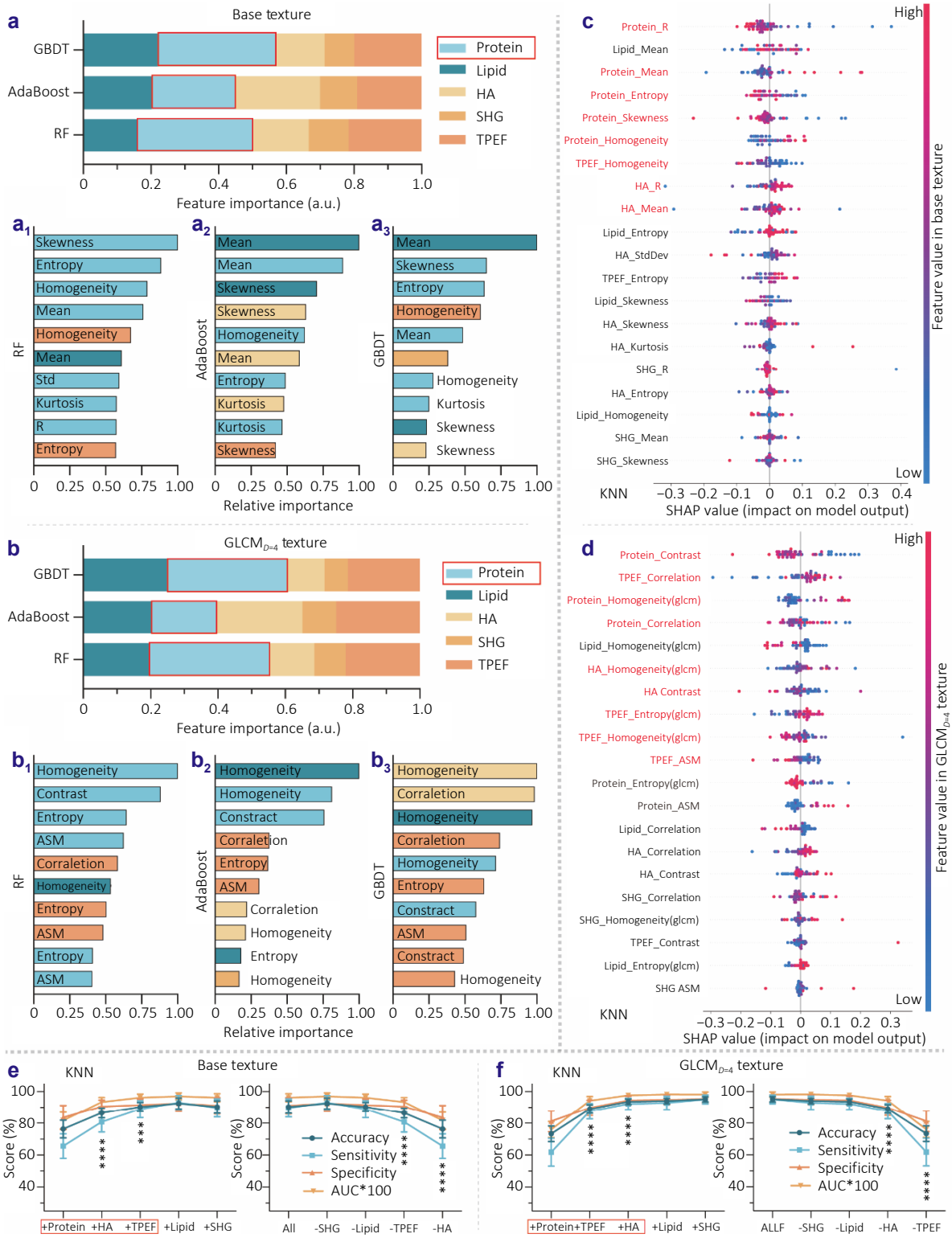

Fig. 5 | Feature importance analysis based on feature fusion. (a, b) Channel-wise contribution percentages in tree-based models (RF, AdaBoost, and GBDT) for base texture (a) and GLCM (D = 4 pixels) features (b). Corresponding panels (a_1–a_3) and (b_1–b_3) display the top 10 critical features for each algorithm. Color coding across all panels is consistent with the channel legend. (c, d) SHAP values of base texture (c) or GLCM (D = 4 pixels) features (d) in KNN model. Feature names are denoted as [Channel]_[Feature] (e.g., Protein_skewness). (e, f) Classification performance changes of KNN model based on fused base texture (e) or GLCM (D =4 pixels) features (f) when adding (left) or removing (right) channel features stepwise. Data were presented as mean with mean ± 95% CI. *p<0.05, **p<0.01, ***p<0.001 by One-way ANOVA with Tukey's *post-hoc* test.

complexity but also reduced experimental workload by identifying essential channels. A greedy algorithm dictated the sequence addition or removal of channels, maximizing AUC improvements at each step. The performance of each channel combination was robustly evaluated by averaging outcomes from 20 random train-test splits. This iterative approach was applied to both base texture and GLCM features (D = 1, 2, 4 pixels) across all models utilized during feature concatenation (Fig. 5(e, f) and Fig. S10).

Consistently across almost all models and feature types, the protein channel was the first to be added or the last to be removed in almost all models and feature types, reconfirming its superior discriminative power as the most differentiated component between the two groups and contributing over 70% classification accuracy. Each channel addition significantly boosted classification accuracy up to three channels, beyond which further additions yielded no substantial improvement. The concatenation of three-channel texture features increased accuracy from approximately 70% (single-channel) to around 85%, consistently stabilizing the AUC value above 0.9. Across various feature types and models, combinations of protein, TPEF, and HA channels, or of protein, TPEF, and lipid channels consistently constituted the highest-performing three-channel ensembles. Notably, for the GLCM (D = 4 pixels) features, the three-channel combination of protein, TPEF, and HA achieved a peak accuracy of 93.56% and an AUC of 0.98 with the KNN classifier (Fig. 5(e)). Analogous trends were observed during sequential channel removal (Fig. 5(e, f) and Fig. S11). These phenomena illustrated that three optimally selected channels are sufficient to robustly characterize most T2DM-related osteon changes. However, it is crucial to acknowledge that the greedy algorithm employed here inherently favors local optima, potentially obscuring global optimal solutions and thus potentially misrepresenting the true contributions of individual features across channels.

To compensate for the shortcomings of this approach, we additionally employed Shapley additive explanation (SHAP) values to explain individual feature contributions (Fig. 5(c) base texture, Fig. 5(d) GLCM (D = 4 pixels), Fig. S12(a) GLCM (D = 1 pixels), Fig. S12(b) GLCM (D = 2 pixels)). The top-preforming KNN classifier was selected to representatively illustrate the SHAP results. SHAP analysis revealed that the R and contrast of the protein channel were the base texture and GLCM (D = 4 pixels) features with the highest contribution to KNN classifier, respectively. The ranking of SHAP values corroborated the findings from Gini feature importance in tree-based models, and further elucidated the relationship between features and model predictive outputs. For base texture, positive contributions to T2DM prediction were primarily associated with the mean and homogeneity of the protein channel, the mean and R of the HA channel, and the entropy of the lipid channel. In contrast, the R, entropy, skewness of protein channel, alongside the homogeneity of TPEF channel were negative to T2DM prediction. For GLCM features, T2DM-related osteons were notably characterized by lower contrast and higher homogeneity in protein distribution, and higher correlation and entropy but lower homogeneity and ASM in the TPEF signal. Collectively, three distinct methods for assessing feature importance consistently highlighted protein as the most effective component for delineating T2DM-related osteon alterations, while also emphasizing the indispensable contribution of other components, especially TPEF.

Channel fusion based on modified stacking ensemble algorithm. A modified stacking ensemble model was additionally employed to achieve channel fusion via a second-level classifier, offering an alternative to feature-level concatenation. This stacking approach demonstrated robustness to specific classifier choices, with most models achieving classification accuracies exceeding 80% and AUC values surpassing 0.85, except when SVM was used as a first-level classifier or a decision tree as a second-level classifier (Fig. S13(a, b)). While these stacking models outperformed single-channel classification, their performance gains were comparatively modest relative to feature concatenation (which yielded accuracies >90%). We further compared the classification performance of different classifiers, finding RF to exhibit slightly superior accuracy among first-layer classifiers, and thus selected it for subsequent comparisons between different texture features (Fig. 6(a), Fig. S13(c)). Intriguingly, the classification accuracy of GLCM features was significantly lower than base texture features in modified stacking models, in sharp contrast with the advantage of GLCM in feature concatenation. The GLCM distance hyperparameter likewise influenced the recognition accuracy in this model fusion context, with D = 2 pixels proving most appropriate.

To precisely quantify the contribution of each channel to final decision, we introduced conditional probability (CP) as a measure of channel importance. CP was defined as the proportion of samples correctly predicted by the first-level model among those correctly predicted by the second-level model. Since each channel corresponded to a first-level model, CP revealed the contribution of each channel in the final prediction. For higher-accuracy models (e.g., those employing RF or KNN as first-level classifiers), the CPs of protein, TPEF and HA channels were notably similar, clustering around 80% (Table S6 base texture, Table S7 GLCM (D = 2 pixels)). In contrast, the CP of protein channel was significantly higher than other channels in lower-accuracy models (e.g., those with SVM(rbf) as the first-level classifier). Lipid and SHG channels exhibited similar CPs across all stacking models, typically ranging from 62% to 73%. Moreover, disregarding model variability, protein consistently exhibited the largest classification contribution across both base texture (Fig. 6(b)) and GLCM (D = 2 pixels, Fig. S13(d)) features, followed by TPEF and HA channels.

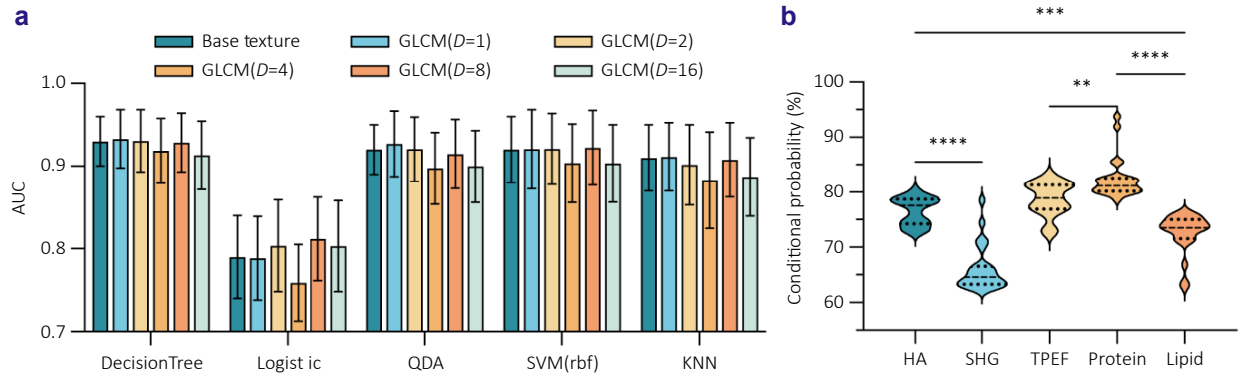

Fig. 6 | Classification performance and channel importance of modified stacking ensemble models. (a) AUC of different features and second-level classifiers. Different features included base texture features and GLCM features at 5 distances (D = 1, 2, 4, 8, 16 pixels), respectively. The first-level classifier was RF model. The second-level classifiers included logistic regression, decision tree, QDA, SVM(rbf), and KNN. Each model was tested 50 times with randomly divided train and test sets. (b) Conditional probability distribution of base texture features under different classifier combinations between different channels. Data were shown as mean with 95% CI. ns p>0.05, *p<0.05, **p<0.01, ***p<0.001 by Two-way ANOVA (a) or One-way ANOVA (b) with Tukey's *post-hoc* test.

2.3 GLCM reveals aberrant multi-component spatial patterns in T2DM osteons

Building upon our classification results from single-channel, feature, and model fusion, we identified GLCM features from protein, TPEF, and HA channels as the most discriminative for characterizing T2DM cortical bone lesions. To visually decipher these T2DM-induced spatial changes in multi-component bone architecture, we mapped the abstract GLCM features to concrete images, utilizing 32 gray levels for visualization (Fig. 7).

For the protein channel, T2DM osteons exhibited GLCM signals predominantly clustered along the matrix diagonal, indicating minimal signal intensity variations within a 4-pixel range. This suggests a more uniform local distribution and gradual changes in protein signals in T2DM bone. Conversely, control GLCM signals were more dispersed around the diagonal, reflecting point-like protein aggregates with sharper signal transitions. These protein distribution changes in T2DM osteons were visually confirmed by NLO images (Fig. 7(a), upper panels) and quantitatively captured by higher homogeneity, correlation, and ASM, alongside lower entropy and contrast in GLCM features (Fig. 7(b), upper panels). In the TPEF channel, due to overall lower signal intensity, GLCM signals were mostly clustered in the upper-left matrix corner. While control GLCM signals clustered within the intensity range [0, 32], T2DM signals extended into the range of [32, 64], consistent with the higher tail observed in its signal histogram (Fig. S5(b)). The GLCM feature of TPEF showed higher correlation and ASM in T2DM. For the HA channel, although both groups displayed diagonal GLCM distributions, control signals clustered in the lower-right (higher intensity) while T2DM signals clustered in the upper-left (lower intensity). Notably, the GLCM discrepancy in HA between groups was considerably less pronounced than in the protein channel, a trend reflected in their respective GLCM features. As anticipated from their limited classification contributions, GLCMs for lipid and SHG channels showed high similarity between groups (Fig. 7, lower panels).

2.4 Higher homogeneity of protein in T2DM osteons

To rigorously compare osteon-level texture features, we averaged the base texture and GLCM features derived from eight sub-images comprising each osteon. This analysis unequivocally revealed that protein channel was the primary differentiator between T2DM and control groups (Fig. 8, Fig. S14, and Fig. S15). Additionally, only the skewness of TPEF channel exhibited significant inter-group differences.

Consistently across both base texture and GLCM feature types, T2DM osteons exhibited increased homogeneity and decreased entropy in protein distribution (Fig. 8). Crucially, similar mean values suggested no significant alteration in the average protein concentration within osteons due to T2DM (Fig. 8(a)). Thus, these findings collectively suggested a reduction in protein aggregation, where proteins originally concentrated in specific locales become more uniformly dispersed over a broader area. This interpretation was further corroborated by GLCM features, with higher ASM and lower contrast directly supporting a more uniform and diffuse protein distribution in T2DM osteons.

3 Discussion

In this study, we employed multimodal label-free NLO microscopy, integrating SRS, SHG, and TPEF, to investigate T2DM-associated alterations in microstructure and material composition of cortical osteons in human lumbar vertebrae. To our knowledge, this is the first application of label-free NLO microscopy to characterize T2DM-specific skeletal

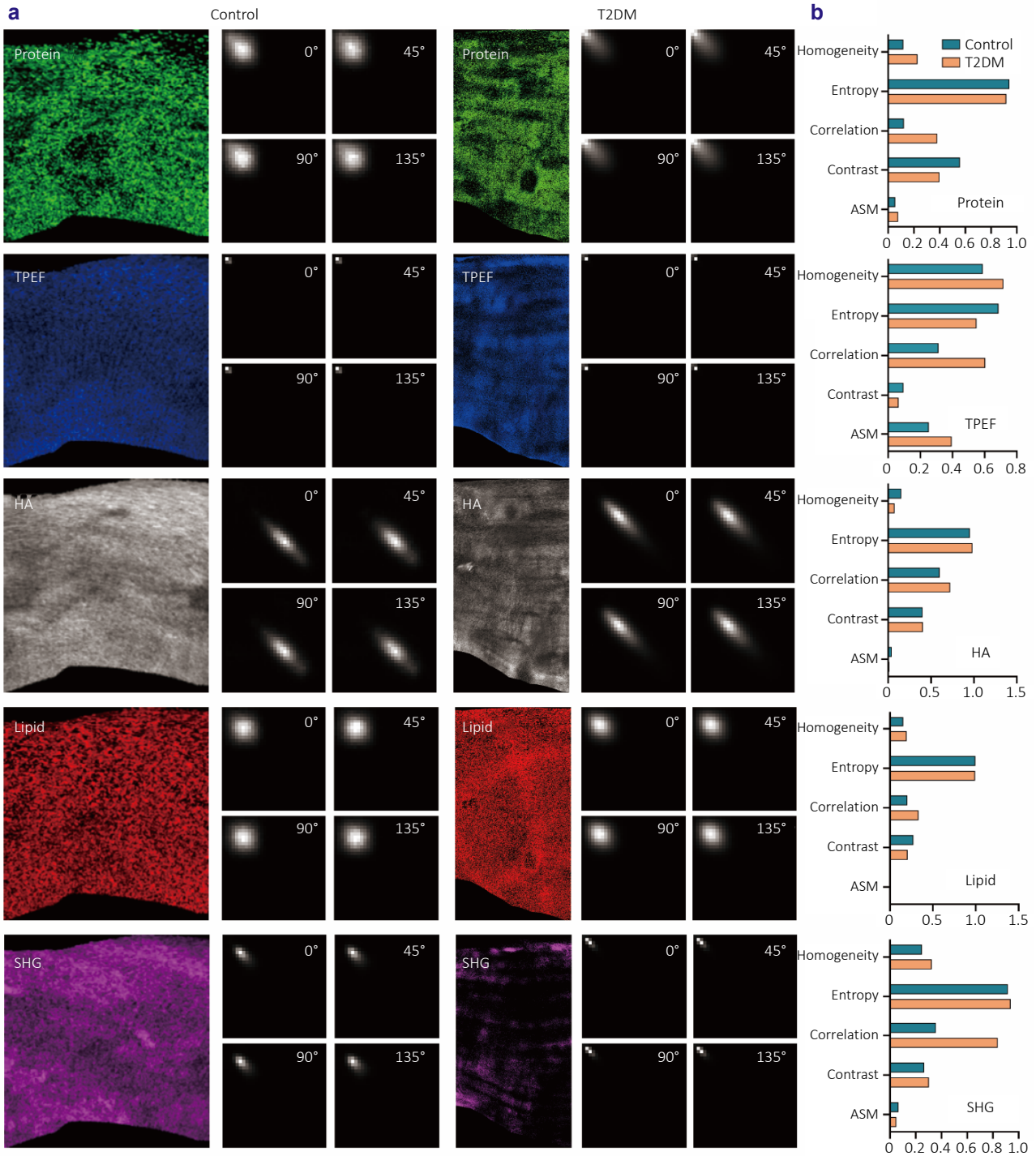

Fig. 7 | GLCM (D = 4 pixels) features from five channels for distinguishing control and T2DM patients. (a) The representative sub-images obtained through the image preprocess step in Fig. 3, and visual grayscale maps of GLCM (D = 4 pixels) in four directions (0°, 45°, 90°, 135°), including protein, TPEF, HA, lipid, and SHG channels in control and T2DM groups. (b) GLCM features averaged over the four directions based on the corresponding channel matrices obtained in (a). The average features were restricted to the range of 0 to 1 by min-max normalization.

alterations in human bone samples. Our primary finding suggested that aberrant spatial distribution of osteocyte-related proteins, characterized by increased homogeneity and decreased contrast, might serve as a distinctive pathological indicator of impaired bone quality observed in T2DM. Multi-channel texture analysis demonstrated strong discriminative power, especially when fusing features from protein, HA, and autofluorescent components. This fused

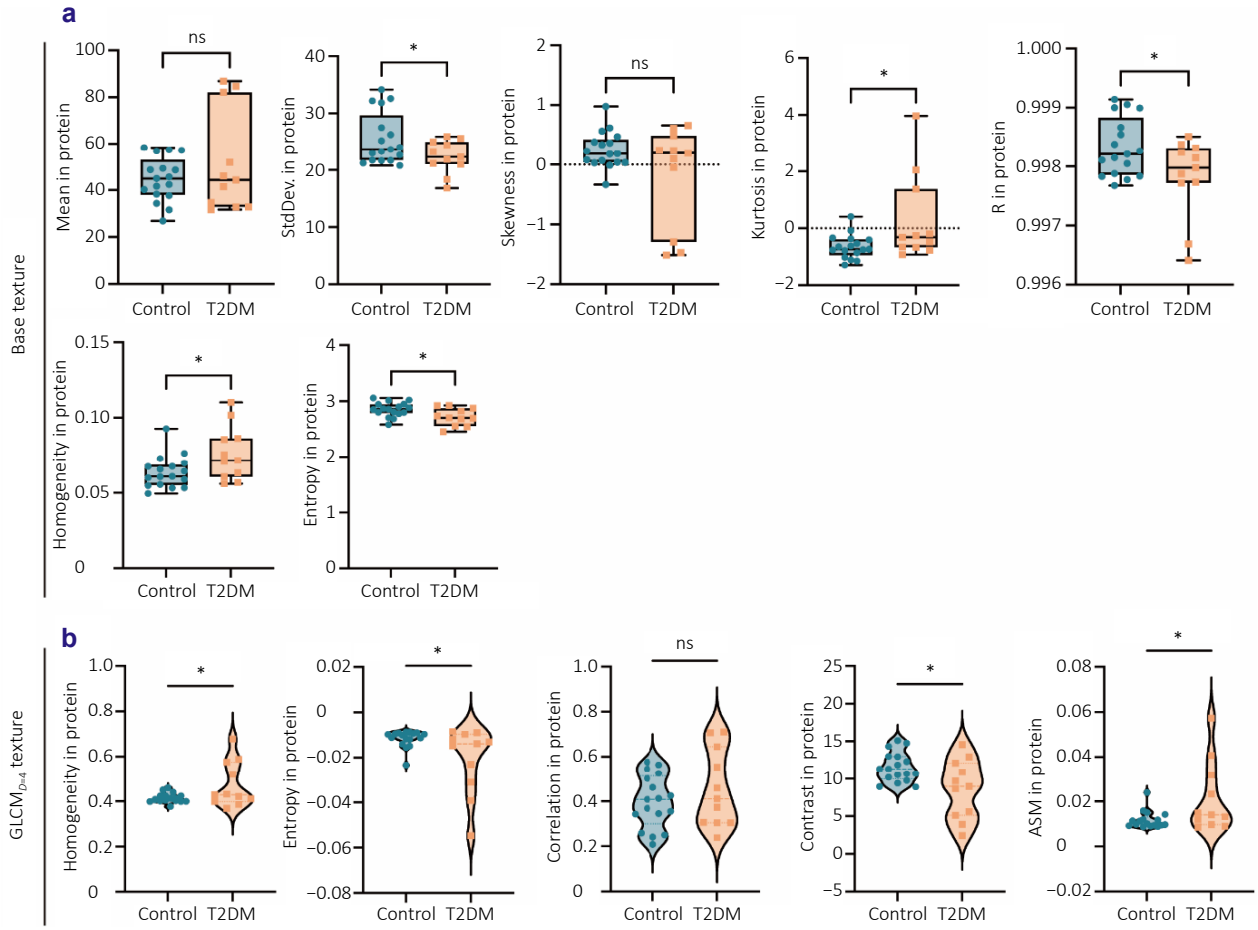

Fig. 8 | Base Texture and GLCM (D = 4 pixels) features from protein channels in control and T2DM osteons. (a) Base texture features. (b) GLCM (D = 4 pixels) features. All P values were calculated using a two-sided Student's t-test. n = 17 in control group, n = 11 in T2DM group. ns p>0.05. *p<0.05.

approach achieved a superior classification performance (93.56% accuracy with KNN classifier) significantly outperforming single-channel analyses (~75%).

The effectiveness and potential of multimodal label-free NLO imaging have been confirmed in numerous critical applications, such as rapid diagnosis[20,22], and high-resolution in vivo imaging[19,59]. Here, we confirmed its suitability for imaging human cortical bone, notably providing both microstructural and multi-material composition information. SRS, SHG and TPEF images clearly visualized key components and architecture of osteons, including lamellae, osteocytes, LCN, and cement line, which was consistent with the results of conventional stain-based imaging[46,60]. Following linear decomposition, SRS images more accurately highlighted the histological features of proteins and lipids, especially those within osteocytes. Although collagen fibers are widely recognized as critical determinants of bone quality, no significant differences were observed in the orientation of collagen fibers via SHG. This can be attributed to the fact that SHG primarily characterizes the quantity and directional alignment of non-centrosymmetric collagen fiber bundles[17,48,61]. Thus, our findings suggest that the overall architectural organization and volume of collagen fibers remain relatively stable in T2DM bone, with pathological alterations likely occurring predominantly at the molecular level (e.g., glycation or advanced glycation end-products). Compared to conventional techniques applied in previous studies to investigate osteons, such as X-ray computed tomography[62], stain-based imaging[60], and others[47,35,63], NLO imaging eliminates the need for decalcification and specific labeling, while delivering high-resolution structural and material information without preprocessing.

The prominence of protein texture as the most influential feature in our model underscores the vulnerability of the osteocyte-related network in T2DM. Rather than a simple reduction in total protein mass, the observed spatial homogenization and decreased contrast suggest a loss of the structural and biochemical complexity required for normal bone homeostasis. In healthy bone, proteins within osteocytes and the LCN are highly organized to support mechanosensing and intercellular communication[64], such as Fibroblast growth factor (FGF-23)[65,66], sclerostin[67,68], connexins, integrins, and others[38,69,70]. The transition toward a more uniform protein distribution in T2DM osteons likely reflects

a dismantling of these functional gradients. While SRS imaging at the 2930 cm^{-1} vibration mode captures a collective signal from various methyl-rich biomolecules rather than specific proteins, these spatial alterations may provide a potential imaging indicator for the dysregulation of the osteocyte mechanotransduction network. Osteocytes contain multiple mechanosensors, such as sclerostin and Connexin 43 (Cx43)[71], which are pivotal for mediating communication both within and between cells. Given that high-glucose environments are known to downregulate Cx43 levels and impair gap junction functions[72,73], it is tempting to speculate that the homogenized protein texture we identified represents a macroscopic manifestation of these microscopic molecular deficits. The observed inverse relationship between intracellular lipid and protein content, characterized by lipid accumulation and protein decrease, further underscores the metabolic disturbances within diabetic osteocytes. This "lipid-high, protein-low" profile suggests a potential shift toward abnormal lipid metabolism at the expense of functional protein synthesis, which may reflect a compromised regulatory capacity of osteocytes in the T2DM environment. Such anomalous lipid accumulation mirrors the pathomorphological changes observed in other models of skeletal fragility, such as alcohol-induced bone loss[74], reinforcing the hypothesis that cellular metabolic reprogramming is a significant contributor to bone quality decline. Our AI-driven identification of this pattern offers a potential pathological indicator, suggesting that osteocyte metabolic dysfunction may be an early-stage contributor to the decline of bone quality in T2DM.

The adoption of texture analysis[75,76] in this study moves beyond simple intensity-based quantification, offering a robust descriptor for the intricate spatial organization of the bone matrix. Texture is extensively applied in image classification and segmentation across various medical imaging modalities, especially in CT[51,53], MRI[52,55], and others[77]. While conventional statistical metrics (base texture) provide a baseline for classification, the significant improvement in accuracy achieved by integrating GLCM features underscores that the primary pathological signature of T2DM osteons resides in the spatial relationship of molecular signals. This suggests that diabetic bone deterioration is characterized by a "spatial disorganization" of the matrix, which traditional intensity measurements fail to capture. Notably, the optimal classification performance observed at GLCM distances of 1, 2, and 4 pixels, corresponding to the physical scale of the bone canalicular diameter[36], provides a biological bridge between our computational features and the bone microstructure. This alignment suggests that our AI model is not merely learning random patterns, but is specifically sensitive to the spatial redistribution of proteins and lipids within LCN. The high discriminative power at these specific scales implies that T2DM disrupts the molecular gradients within the LCN, leading to the "homogenized" appearance identified in our texture maps. By quantifying these spatial link properties, our approach effectively translates abstract imaging textures into tangible indicators of LCN integrity, further corroborating the qualitative evidence of metabolic and structural dysfunction in diabetic osteocytes.

The superior performance of the KNN algorithm in the early fusion stage (feature concatenation) can be attributed to its non-parametric, distance-based nature, which contributed to capturing local manifold structures and clustering patterns without assuming a global data distribution. In our study, multimodal features concatenation produced a high-dimensional feature space in which discriminative information emerged primarily from joint cross-channel patterns rather than from any single modality. Thus, under the constraints of our current pilot-scale sample size, KNN's reliance on local "neighborhood" consistency allowed it to capture these subtle structures with high sensitivity. In contrast, ensemble methods (e.g., RF) rely on stochastic feature sub-sampling during decision tree construction. Although this mechanism enhances generalization in many applications, it may reduce the algorithm's ability to exploit multimodal interaction information. In particular, RF's feature randomization might dilute subtle inter-channel dependencies that KNN's local manifold preservation captures effectively in this dataset. However, it is critical to note that as the dataset expands, KNN may encounter the "curse of dimensionality"[78]. This phenomenon diminishes the discriminative power of distance metrics in high-dimensional spaces, potentially increasing sensitivity to noise. RF demonstrated stronger performance in the stacking stage, where its ensemble architecture proved robust in aggregating non-linear features. As data volume grows, the ensemble learning and feature-randomization mechanisms of RF are expected to offer superior generalizability and robustness, making it better equipped to handle the increased biological variability and complexity of larger populations than distance-based classifiers.

Some limitations should be acknowledged in this study. First, the relatively limited patient-level size may restrict the generalizability of our findings to the broader and highly heterogeneous T2DM population. Although we utilized a rigorous patient-wise cross-validation framework with 50 random partitions to enhance robustness and found no significant baseline differences, future studies with larger, multi-center cohorts are essential to mitigate the influence of biological variability and other clinical confounders (e.g., diabetes duration, medications). Second, the observed protein spatial homogenization identified in this study should be cautiously interpreted as an imaging-based pathological indicator rather than a confirmed molecular mechanism. SRS signals at the 2930 cm^{-1} band reflect collective contributions from methyl-rich biomolecules rather than specific protein species, rendering our biological interpretations (e.g., links to mechanosensing proteins) hypothetical. Orthogonal validations such as immunohistochemistry,

proteomics, or biomechanical testing, are essential to elucidate causal relationships between these texture changes and impaired bone quality. Finally, while our AI-powered NLO workflow effectively discriminates T2DM-associated bone alterations in this pilot study, it currently serves as a discovery tool rather than a fully automated clinical diagnostic platform. Ongoing work will focus on dataset expansion to refine lesion patterns and pursue targeted molecular investigations, advancing from indicator identification toward a deeper understanding of mechanistic underpinnings of diabetic bone disease.

In summary, our findings revealed that aberrant protein texture, characterized by increased homogeneity and decreased contrast, is a critical factor contributing to impaired bone quality in T2DM. A robust workflow based on a KNN model was also developed to integrate texture features from protein, TPEF, and HA channels, demonstrating high classification performance (accuracy 93.56%) in distinguishing T2DM-related skeletal alterations. Collectively, this study not only constructed a comprehensive framework for quantitatively identifying T2DM-related skeletal alterations with high accuracy, but also provided novel insights into the possible pathogenic mechanisms underlying bone microarchitecture alterations and subsequent bone quality deterioration in T2DM patients.

4 Materials and methods

4.1 Setting

In this study, 55 individuals participated from March 2021 to January 2022. Nineteen patients were excluded due to insufficient bone tissue volume for effective sectioning. Consequently, 36 patients were ultimately included in our study. Analysis of bright field images revealed 28 osteons in tissue sections from 22 participants (12 without T2DM, including 6 males, aged 63.3±15.0 years; and 10 T2DM, including 5 males, aged 68.9±12.0 years). This study was approved by the local ethics committee of Shanghai East Hospital (EC.D (BG) 016. 02.1). Each participant had signed an informed consent form. Before the operation, the demographic and biochemical profiles of the T2DM and control groups were comparable, with the notable exceptions of glycemia and HbA1c levels (Table S1). Since all participants had lumbar spinal stenosis, cortical tissues of L4 (or L5) were collected when they underwent posterior approach spinal decompression (discectomy and laminectomy). Skeleton samples were cut into thin slices using a rotary microtome before optical microscopy (Fig. S1(a)). The thickness of each slice was around 100 μm. Then they were placed on slides with coverslips and nail polish was applied to the outer for encapsulation.

4.2 Nonlinear optical microscopy

Detailed setup of the multimodal NLO microscopy has been illustrated in previous studies[21,44,79]. In brief, we employed a picosecond pulse laser (picoEmerald S, Applied Physics & Electronics) with an 8 MHz repetition rate and 2 ps pulse width, which integrated the output for the pump beam (Tunable wavelength, 700~960 nm) and the Stokes beam (Fix wavelength, 1031 nm). In SRS imaging, the Stokes beam was modulated at approximately 20 MHz using an electro-optic modulator. The collinear pump and Stokes beams were coupled to a two-dimensional scanning Galvanometer mirror system (GVS012-2D, Thorlabs) and introduced into an inverted research microscope (model IX73, Olympus). Two 60× water-immersion objective lenses (LUMPlanFL N, 1.0 numerical aperture, Olympus) were used to illuminate the sample and collect the transmission of forward-detected SRS signal, respectively. The detected signal was then processed by a lock-in amplifier (model HF2LI, Zurich Instruments). To detect the vibrational bands of CH_2 (2850 cm^{-1}), CH_3 (2930 cm^{-1}), and PO_4^{-3} (959 cm^{-1}), the pump beam wavelength was adjusted to 796.8 nm, 791.8 nm, and 938 nm, respectively. The excitation power at the sample was maintained at approximately 20 mW for the pump beam and approximately 90 mW for the Stokes beam across all bone tissue samples. The same microscope was used to collect both SHG and TPEF imaging. The pump beam was tuned to 796.8 nm, and the excitation power at the sample was ~40 mW. A photomultiplier tube (H7422-40, Hamamatsu) was used to detect the backward SHG and TPEF signals through 395/25 nm (ET395/25, Chroma) and 520/40 nm (ET520/40, Chroma), respectively. All NLO images were obtained in 512 × 512 pixels with dwell time of 8 μs. High-resolution microscopic images of tissue sections were acquired using a motorized stage microscope (model H117, Prior Scientific). No photodamage to the tissue samples was observed during the imaging process.

4.3 Osteons separation

Secondary osteons (excluding haversian canals) were selected as the primary units for comparison to ensure structural consistency (Fig. S3 for the images with and without haversian system). SHG, TPEF, and SRS-HA channels were utilized to differentiate osteons from the surrounding matrix. To ensure segmentation accuracy and minimize subjective bias, we employed a hybrid workflow involving algorithmic drafting and multi-observer refinement using ImageJ (https://imagej.nih.gov/ij/): 1) **Algorithmic drafting**: K-means clustering was first applied to the TPEF channel to generate a preliminary segmentation pseudo-color map (Fig. S4(b)). This algorithmic step effectively differentiated the circumferential lamella within the osteon from the interstitial lamellae based on signal intensity gradients. 2) **Manual adjustment**: preliminary boundaries were initially drafted

and then independently refined by two trained observers using the "polygon selections" tool. This step ensured that the complex morphology of secondary osteons was captured accurately across all channels. 3) **Expert adjudication**: a third experienced researcher performed the final adjudication of all segmented boundaries. The interior and exterior ROIs were merged via the "XOR" function in the ROI Manager to isolate the final osteon area. Identical regions across all channels were then extracted and saved as the maximum bounding rectangles for further analysis (Fig. S4).

4.4 Linear decomposition of SRS

The measured spectra image was represented as a linear combination of Raman spectra of standard substances. The signals at the Raman shift of 2930 cm^{-1} and 2850 cm^{-1} were both mainly composed of protein and lipid components. Bovine serum albumin (BSA) and glyceryl trioleate (TO) were used as the protein and lipid standards, respectively. The coefficient matrix in Eq. (1) was obtained from the normalized spontaneous Raman spectra of BSA and TO (Fig. S1(b)), by extracting their intensities at 2930 cm^{-1} and 2850 cm^{-1}. These values define the linear mixing coefficients between the two channels and the two biochemical components. By solving Eq. (1) using this system-specific calibration matrix, the decomposed SRS images of protein and lipid channels could be obtained.

$$\begin{pmatrix} X_{2930\ cm^{-1}} \\ X_{2850\ cm^{-1}} \end{pmatrix} = \begin{pmatrix} 0.85 & 1 \\ 0.93 & 0.18 \end{pmatrix} \begin{pmatrix} X_{\text{lipid}} \\ X_{\text{protein}} \end{pmatrix}. \tag{1}$$

4.5 Orientation measurement of collagen fibers

We employed the plugin of ImageJ (OrientationJ, http://bigwww.epfl.ch/demo/orientation/) to measure the orientation distribution in the separated SHG images of osteons. Firstly, the Gaussian filter (kernel size: 3×3, μ=0, σ=1) was utilized for removing noise signals. Secondly, each image was divided into numerous 40×40 pixel blocks. The local orientation, energy, and coherency of collagen fibers were calculated in each block by "OrientationJ Vector Field". The orientation range was set from −90° to 90°. Finally, the energy and coherency of one osteon was the average value of constituent blocks. The orientation of each osteon was manifested through the distribution of orientations in blocks.

4.6 Multi-channel quantitative statistics

Extracted images in each channel were processed by the Gaussian filter (kernel size: 3×3, μ=0, σ=1). Then, the valid signal region was selected and measured using the "Threshold" and "Measure" functions in ImageJ respectively. The statistical indicators included mean, standard deviation (StdDev), skewness, and kurtosis.

4.7 Image preprocessing

Firstly, the extracted osteon images were processed by the Gaussian filter (kernel size: 3×3, μ=0, σ=1) to suppress high-frequency noise and ensure the reliability of feature extraction. Then the processed images were converted into polar coordinates. The center of transformation was the central point of the Haversian canal. Secondly, each transformed image was divided into 8 sub-images in accordance with equal width, and the empty pixel blocks at the top and bottom edges were removed. It was equivalent to dividing one sub-image per 45° in raw images. The sub-images were the basic units for feature extraction and classification. Each channel had 224 sub-images (88 from T2DM, 136 from control). Finally, each sub-image in the training set was augmented twice to expand the dataset. Notably, all images generated from data augmentation were only used in the training set, and all images in the test set were actual images. The augmentation techniques included flipping (up and down, side to side), rotation (90°, 180°, 270°), transposing, adding Gaussian noise or pepper noise, and resizing (height to half). We randomly generated two numbers to choose two augmentation techniques for each sub-image. Given the naturally annular lamellar structure and inherent rotational invariance of osteons, as well as the orientation independence of all extracted features, these image augmentation methods do not distort the intrinsic microstructural information of osteons, ensuring high reliability.

4.8 Texture feature extraction

We extracted two types of texture features from each sub-image in five channels, including base texture features and GLCM features[80,81]. Base texture features were statistical features based on signal intensity distribution, including mean, StdDev, skewness, kurtosis, *R*, homogeneity, and entropy. GLCM features were based on the grayscale co-occurrence pattern of adjacent pixels, including homogeneity, correlation, contrast, entropy, and angular second moment (ASM)[57]. The detailed calculation formulas were listed in supplementary materials. The GLCM features involved three hyperparameters: gray levels, distances, and directions. We evaluated 6 gray levels (L = 256, 128, 64, 32, 16, 8), along with 5 distances (D = 1, 2, 4, 8, 16 pixels) and 4 directions (θ = 0°, 45°, 90°, 135°)[58]. Lower gray levels simplified computational complexity but risked image distortion. Thus, we compared 6 gray levels of sub-images with five channels (Fig. S6) and found that the 32 gray-level optimally balanced algorithmic complexity and feature accuracy, which was adopted for further analysis. Furthermore, the GLCM matrices of four directions exhibited a high degree of similarity, prompting the computation of the average GLCM vector of four directions as the representative GLCM feature at the corresponding distance. However, due to the lack of consensus on the most effective distance parameters for

osteon classification, we extracted the 4-direction average GLCM features across the 5 distances.

4.9 Model building for classification

Single channel classification. Eight models were employed for single-channel classification, including SVM(rbf), Decision Tree, RF, Naïve Bayesian, Logistic Regression, AdaBoost, GBDT, and KNN. The texture features extracted from individual channels served as the input for the models.

Channel fusion based on concatenating features. We concatenated features extracted from five channels to form a higher-dimensional feature set, which served as input to the ML algorithms. In this stage, SVM(rbf), RF, AdaBoost, GBDT, and KNN models were selected to distinguish osteons between the control and T2DM groups.

Channel fusion based on modified stacking ensemble algorithm. The modified stacking algorithm consisted of two layers of classifiers. The first layer comprised 5 identical fundamental classifiers, each corresponding to the texture features of one distinct channel. The second layer, as a decision classifier, achieved channel-fusion prediction by using the predicted probability matrix output from the first-layer classifiers. We experimented with five models used in feature concatenation classification as the first-layer classifiers, including RF, SVM(rbf), KNN, GBDT, and AdaBoost, and selected lower complexity models as the second-layer algorithms, including logistic regression, decision tree, quadratic discriminant analysis (QDA), SVM(rbf), and KNN. Thus, eight stacking models were built and utilized for classification.

4.10 Training and prediction

For each model, the datasets were randomly split into 75% training sets and 25% test sets based on patients. In total, for each training and prediction, 56 samples were in the test set, and 504 samples were in the training set. 10-fold cross-validation was utilized to evaluate the performance of classification models. In Stages 3–5, to mitigate potential biases introduced by the test set selection on the outcomes, we randomly divided each feature dataset into training and test sets 50 times, generating 50 distinct test sets along with their corresponding training sets. Each model was then trained and tested 50 times using these different train and test sets. The final classification performance of each model was characterized by averaging the results across all 50 iterations. The sequence of random seeds remained constant across the 50 divisions of each feature set to ensure comparability between features and algorithms. The main evaluation indicators of model prediction performance included accuracy, sensitivity, and specificity. Receiver Operating Characteristic (ROC) curves were also plotted and AUC was calculated as the comprehensive evaluation indicator.

4.11 Statistical analysis

The statistical analysis in our study was conducted and illustrated using the Prism software (GraphPad, version 10, https://www.graphpad.com/scientific-software/prism/). The quantitative results were described as mean ± 95% CI. The independent samples *t*-test was used to examine the differences in indicators of normal distribution between T2DM and control groups. Welch correction was utilized for the indicators that did not conform to homogeneity of variance. Mann-Whitney test was used for comparing indicators of abnormal distribution. To compare multiple groups, one-way or two-way analysis of variance (ANOVA) was used with Tukey's multiple-comparison *post hoc* test. Two-tailed $p < 0.05$ was regarded as the significant level for all tests.

References

1. GBD 2021 Diabetes Collaborators. Global, regional, and national burden of diabetes from 1990 to 2021, with projections of prevalence to 2050: a systematic analysis for the Global Burden of Disease Study 2021. *Lancet* **402**, 203–234 (2023).
2. Koromani F, Oei L, Shevroja E et al. Vertebral fractures in individuals with type 2 diabetes: more than skeletal complications alone. *Diabetes Care* **43**, 137–144 (2020).
3. Zimmet PZ, Magliano DJ, Herman WH et al. Diabetes: a 21st century challenge. *Lancet Diabetes Endocrinol* **2**, 56–64 (2014).
4. Shanbhogue VV, Hansen S, Frost M et al. Bone disease in diabetes: another manifestation of microvascular disease. *Lancet Diabetes Endocrinol* **5**, 827–838 (2017).
5. Hofbauer LC, Busse B, Eastell R et al. Bone fragility in diabetes: novel concepts and clinical implications. *Lancet Diabetes Endocrinol* **10**, 207–220 (2022).
6. Hofbauer LC, Brueck CC, Singh SK et al. Osteoporosis in patients with diabetes mellitus. *J Bone Miner Res* **22**, 1317–1328 (2007).
7. Khosla S, Samakkarnthai P, Monroe DG et al. Update on the pathogenesis and treatment of skeletal fragility in type 2 diabetes mellitus. *Nat Rev Endocrinol* **17**, 685–697 (2021).
8. Leslie WD, Rubin MR, Schwartz AV et al. Type 2 diabetes and bone. *J Bone Miner Res* **27**, 2231–2237 (2012).
9. Shanbhogue VV, Mitchell DM, Rosen CJ et al. Type 2 diabetes and the skeleton: new insights into sweet bones. *Lancet Diabetes Endocrinol* **4**, 159–173 (2016).
10. Seeman E, Delmas PD. Bone quality — the material and structural basis of bone strength and fragility. *N Engl J Med* **354**, 2250–2261 (2006).
11. Tokarz D, Martins JS, Petit ET et al. Hormonal regulation of osteocyte perilacunar and canalicular remodeling in the hyp mouse model of X-linked hypophosphatemia. *J Bone Miner Res* **33**, 499–509 (2018).
12. Pitfield R, Deter C, Mahoney P. Bone histomorphometric measures of physical activity in children from medieval England. *Am J Phys Anthropol* **169**, 730–746 (2019).
13. Sugawara Y, Kamioka H, Honjo T et al. Three-dimensional reconstruction of chick calvarial osteocytes and their cell processes using confocal microscopy. *Bone* **36**, 877–883 (2005).
14. Mabilleau G, Perrot R, Flatt PR et al. High fat-fed diabetic mice present with profound alterations of the osteocyte network. *Bone* **90**, 99–106 (2016).
15. Yue S, Slipchenko MN, Cheng JX. Multimodal nonlinear optical microscopy. *Laser Photon Rev* **5**, 496–512 (2011).
16. Fu D, Lu FK, Zhang X et al. Quantitative chemical imaging with multiplex stimulated Raman scattering microscopy. *J Am Chem Soc*

134, 3623–3626 (2012).
17. Pendleton EG, Tehrani KF, Barrow RP et al. Second harmonic generation characterization of collagen in whole bone. *Biomed Opt Express* **11**, 4379–4396 (2020).
18. Thomas G, van Voskuilen J, Gerritsen HC et al. Advances and challenges in label-free nonlinear optical imaging using two-photon excitation fluorescence and second harmonic generation for cancer research. *J Photochem Photobiol B* **141**, 128–138 (2014).
19. Lu FK, Basu S, Igras V et al. Label-free DNA imaging in vivo with stimulated Raman scattering microscopy. *Proc Natl Acad Sci USA* **112**, 11624–11629 (2015).
20. Ji MB, Arbel M, Zhang LL et al. Label-free imaging of amyloid plaques in Alzheimer's disease with stimulated Raman scattering microscopy. *Sci Adv* **4**, eaat7715 (2018).
21. Jia H, Liu J, Fang TH et al. The role of altered lipid composition and distribution in liver fibrosis revealed by multimodal nonlinear optical microscopy. *Sci Adv* **9**, eabq2937 (2023).
22. Ji MB, Orringer DA, Freudiger CW et al. Rapid, label-free detection of brain tumors with stimulated Raman scattering microscopy. *Sci Transl Med* **5**, 201ra119 (2013).
23. Tang TT, Ebacher V, Cripton P et al. Shear deformation and fracture of human cortical bone. *Bone* **71**, 25–35 (2015).
24. Houle MA, Couture CA, Bancelin S et al. Analysis of forward and backward Second Harmonic Generation images to probe the nanoscale structure of collagen within bone and cartilage. *J Biophotonics* **8**, 993–1001 (2015).
25. Osorio R, Asady S, Toledano-Osorio M et al. Biomimetic remineralization of an extracellular matrix collagen membrane for bone regeneration. *Polymers* **14**, 3274 (2022).
26. Shin KS, Men SQ, Wong A et al. Quantitative chemical imaging of bone tissue for intraoperative and diagnostic applications. *Anal Chem* **94**, 3791–3799 (2022).
27. Zheng W, Li D, Li SX et al. Diagnostic value of nonlinear optical signals from collagen matrix in the detection of epithelial precancer. *Opt Lett* **36**, 3620–3622 (2011).
28. Orringer DA, Pandian B, Niknafs YS et al. Rapid intraoperative histology of unprocessed surgical specimens via fibre-laser-based stimulated Raman scattering microscopy. *Nat Biomed Eng* **1**, 0027 (2017).
29. Wu WT, Brandt C, Zhou X et al. Label-free multimodal imaging with simultaneous two-photon and three-photon microscopy and kernel-based nonlinear scaling denoising. *Biomed Opt Express* **15**, 114–130 (2024).
30. Talone B, Bresci A, Manetti F et al. Label-free multimodal nonlinear optical microscopy reveals features of bone composition in pathophysiological conditions. *Front Bioeng Biotechnol* **10**, 1042680 (2022).
31. Kumar S, Kamali T, Levitte JM et al. Single-pulse CARS based multimodal nonlinear optical microscope for bioimaging. *Opt Express* **23**, 13082–13098 (2015).
32. Chen YG, Liu SJ, Liu HJ et al. Coherent Raman scattering unravelling mechanisms underlying skull optical clearing for through-skull brain imaging. *Anal Chem* **91**, 9371–9375 (2019).
33. Inaba M, Okuno S, Kumeda Y et al. Increased incidence of vertebral fracture in older female hemodialyzed patients with type 2 diabetes mellitus. *Calcif Tissue Int* **76**, 256–260 (2005).
34. Chang B, Liu XH. Osteon: structure, turnover, and regeneration. *Tissue Eng Part B Rev* **28**, 261–278 (2022).
35. Maggiano IS, Maggiano CM, Clement JG et al. Three-dimensional reconstruction of Haversian systems in human cortical bone using synchrotron radiation-based micro-CT: morphology and quantification of branching and transverse connections across age. *J Anat* **228**, 719–732 (2016).
36. Zimmermann EA, Busse B, Ritchie RO. The fracture mechanics of human bone: influence of disease and treatment. *Bonekey Rep* **4**, 743 (2015).
37. Zedda M, Lepore G, Biggio GP et al. Morphology, morphometry and spatial distribution of secondary osteons in equine femur. *Anat Histol Embryol* **44**, 328–332 (2015).
38. Bonewald LF. The amazing osteocyte. *J Bone Miner Res* **26**, 229–238 (2011).
39. Buenzli PR, Sims NA. Quantifying the osteocyte network in the human skeleton. *Bone* **75**, 144–150 (2015).
40. Busse B, Bale HA, Zimmermann EA et al. Vitamin D deficiency induces early signs of aging in human bone, increasing the risk of fracture. *Sci Transl Med* **5**, 193ra88 (2013).
41. Zimmermann EA, Schaible E, Bale H et al. Age-related changes in the plasticity and toughness of human cortical bone at multiple length scales. *Proc Natl Acad Sci USA* **108**, 14416–14421 (2011).
42. Dallas SL, Moore DS. Using confocal imaging approaches to understand the structure and function of osteocytes and the lacuno-canalicular network. *Bone* **138**, 115463 (2020).
43. Gortázar AR, Ardura JA. Osteocytes and diabetes: altered function of diabetic osteocytes. *Curr Osteoporos Rep* **18**, 796–802 (2020).
44. Yan S, Cui SS, Ke K et al. Hyperspectral stimulated Raman scattering microscopy unravels aberrant accumulation of saturated fat in human liver cancer. *Anal Chem* **90**, 6362–6366 (2018).
45. Zhang LY, Shi LY, Shen YH et al. Spectral tracing of deuterium for imaging glucose metabolism. *Nat Biomed Eng* **3**, 402–413 (2019).
46. Dong XN, Qin A, Xu JK et al. In situ accumulation of advanced glycation endproducts (AGEs) in bone matrix and its correlation with osteoclastic bone resorption. *Bone* **49**, 174–183 (2011).
47. Stockhausen KE, Qwamizadeh M, Wölfel EM et al. Collagen fiber orientation is coupled with specific nano-compositional patterns in *Dark* and *Bright* osteons modulating their biomechanical properties. *ACS Nano* **15**, 455–467 (2021).
48. Chen XY, Nadiarynkh O, Plotnikov S et al. Second harmonic generation microscopy for quantitative analysis of collagen fibrillar structure. *Nat Protoc* **7**, 654–669 (2012).
49. Rezakhaniha R, Agianniotis A, Schrauwen JTC et al. Experimental investigation of collagen waviness and orientation in the arterial adventitia using confocal laser scanning microscopy. *Biomech Model Mechanobiol* **11**, 461–473 (2012).
50. Neher P, Hirjak D, Maier-Hein K. Radiomic tractometry reveals tract-specific imaging biomarkers in white matter. *Nat Commun* **15**, 303 (2024).
51. Qiu JJ, Yin J, Qian W et al. A novel multiresolution-statistical texture analysis architecture: radiomics-aided diagnosis of PDAC based on plain CT images. *IEEE Trans Med Imaging* **40**, 12–25 (2021).
52. Eun NL, Kang D, Son EJ et al. Texture analysis with 3.0-T MRI for association of response to neoadjuvant chemotherapy in breast cancer. *Radiology* **294**, 31–41 (2020).
53. Lubner MG, Smith AD, Sandrasegaran K et al. CT texture analysis: definitions, applications, biologic correlates, and challenges. *Radiographics* **37**, 1483–1503 (2017).
54. Zhang J, Yu CS, Jiang GL et al. 3D texture analysis on MRI images of Alzheimer's disease. *Brain Imaging Behav* **6**, 61–69 (2012).
55. Zacharaki EI, Wang SM, Chawla S et al. Classification of brain tumor type and grade using MRI texture and shape in a machine learning scheme. *Magn Reson Med* **62**, 1609–1618 (2009).
56. Castro-Zunti R, Park EH, Choi Y et al. Early detection of ankylosing spondylitis using texture features and statistical machine learning, and deep learning, with some patient age analysis. *Comput Med Imaging Graph* **82**, 101718 (2020).
57. Yang MC, Moon WK, Wang YCF et al. Robust texture analysis using multi-resolution gray-scale invariant features for breast sonographic tumor diagnosis. *IEEE Trans Med Imaging* **32**, 2262–2273 (2013).
58. Torheim T, Malinen E, Kvaal K et al. Classification of dynamic contrast enhanced MR images of cervical cancers using texture analysis and support vector machines. *IEEE Trans Med Imaging* **33**, 1648–1656 (2014).
59. Ozeki Y, Umemura W, Otsuka Y et al. High-speed molecular spec-

tral imaging of tissue with stimulated Raman scattering. *Nat Photon* **6**, 845–851 (2012).
60. Liu TJ, Wang J, Xie XD et al. *DMP1* ablation in the rabbit results in mineralization defects and abnormalities in Haversian canal/osteon microarchitecture. *J Bone Miner Res* **34**, 1115–1128 (2019).
61. Hovhannisyan VA, Hu PS, Tan HY et al. Spatial orientation mapping of fibers using polarization-sensitive second harmonic generation microscopy. *J Biophoton* **5**, 768–776 (2012).
62. Bigley RF, Griffin LV, Christensen L et al. Osteon interfacial strength and histomorphometry of equine cortical bone. *J Biomech* **39**, 1629–1640 (2006).
63. Pazzaglia UE, Congiu T, Marchese M et al. Morphometry and patterns of lamellar bone in human Haversian systems. *Anat Rec* (*Hoboken*) **295**, 1421–1429 (2012).
64. Robling AG, Bonewald LF. The osteocyte: new insights. *Annu Rev Physiol* **82**, 485–506 (2020).
65. Quarles LD. Skeletal secretion of FGF-23 regulates phosphate and vitamin D metabolism. *Nat Rev Endocrinol* **8**, 276–286 (2012).
66. Rhee Y, Bivi N, Farrow E et al. Parathyroid hormone receptor signaling in osteocytes increases the expression of fibroblast growth factor-23 in vitro and in vivo. *Bone* **49**, 636–643 (2011).
67. Van Bezooijen RL, Roelen BAJ, Visser A et al. Sclerostin is an osteocyte-expressed negative regulator of bone formation, but not a classical BMP antagonist. *J Exp Med* **199**, 805–814 (2004).
68. Winkler DG, Sutherland MK, Geoghegan JC et al. Osteocyte control of bone formation via sclerostin, a novel BMP antagonist. *EMBO J* **22**, 6267–6276 (2003).
69. Qin C, D'Souza R, Feng JQ. Dentin matrix protein 1 (DMP1): new and important roles for biomineralization and phosphate homeostasis. *J Dent Res* **86**, 1134–1141 (2007).
70. Kalajzic I, Braut A, Guo D et al. Dentin matrix protein 1 expression during osteoblastic differentiation, generation of an osteocyte GFP-transgene. *Bone* **35**, 74–82 (2004).
71. Zhang DM, Li X, Pi CX et al. Osteoporosis-decreased extracellular matrix stiffness impairs connexin 43-mediated gap junction intercellular communication in osteocytes. *Acta Biochim Biophys Sin* **52**, 517–526 (2020).
72. Yang L, Zhou GP, Li MY et al. High glucose downregulates connexin 43 expression and its gap junction and hemichannel function in osteocyte-like MLO-Y4 cells through activation of the p38MAPK/ERK signal pathway. *Diabetes Metab Syndr Obes* **13**, 545–557 (2020).
73. Li AF, Sato T, Haimovici R et al. High glucose alters connexin 43 expression and gap junction intercellular communication activity in retinal pericytes. *Invest Ophthalmol Vis Sci* **44**, 5376–5382 (2003).
74. Maurel DB, Pallu S, Jaffré C et al. Osteocyte apoptosis and lipid infiltration as mechanisms of alcohol-induced bone loss. *Alcohol Alcohol* **47**, 413–422 (2012).
75. Tuceryan M, Jain AK. Texture analysis. In Chen CH, Pau LF, Wang PSP. *Handbook of Pattern Recognition and Computer Vision* 235–276 (World Scientific, Singapore, 1993). https://doi.org/10.1142/9789814343138_0010.
76. Bharati MH, Liu JJ, MacGregor JF. Image texture analysis: methods and comparisons. *Chemometr Intell Lab Syst* **72**, 57–71 (2004).
77. Legesse FB, Medyukhina A, Heuke S et al. Texture analysis and classification in coherent anti-Stokes Raman scattering (CARS) microscopy images for automated detection of skin cancer. *Comput Med Imaging Graph* **43**, 36–43 (2015).
78. Kouiroukidis N, Evangelidis G. The effects of dimensionality curse in high dimensional kNN search. In *2011 15th Panhellenic Conference on Informatics* 41–45 (IEEE, 2011). http://doi.org/10.1109/PCI.2011.45.
79. Zhou WH, Liu DN, Fang TH et al. Rapid and precise diagnosis of retroperitoneal liposarcoma with deep-learned label-free molecular microscopy. *Anal Chem* **96**, 9353–9361 (2024).
80. Park S, Kim B, Lee J et al. GGO nodule volume-preserving nonrigid lung registration using GLCM texture analysis. *IEEE Trans Biomed Eng* **58**, 2885–2894 (2011).
81. Gomez W, Pereira WCA, Infantosi AFC. Analysis of Co-occurrence texture statistics as a function of gray-level quantization for classifying breast ultrasound. *IEEE Trans Med Imaging* **31**, 1889–1899 (2012).

Acknowledgements

This research was funded by National Key R&D Program of China (No. 2025YFE0204500, 2025YFE0218400, 2025YFC2427800, 2025ZD0548504), the Major Research plan of the National Natural Science Foundation of China (No. 92570204), Tianjin Municipal Civil–Military Integration Development Demonstration and Guidance Fund—Project (No. 202309), and PUMC Outstanding Talent Program (2025-I2M-XHJC-045). We sincerely thank Prof. Felix Wang (KU Leuven), Prof. Biao Sun (Tianjin University) for editing and polishing the paper.

Author contributions

B. Z. and J. P. performed all analyses, data curation, validation, and software. S. Y., H. Z., and J. Z. preprocessed the tissues and collected all raw images. T. H. collected the tissues from surgery operations. B. Z. wrote the paper. Z. C. and X. J. assisted with data analysis. L. L. revised the paper. T. L. conceptualized this research, proposed the methodology, revised the paper, provided funding support, and supervised the work.

Competing interests

The authors declare no competing financial interests.

Data availability

Data underlying the results presented in this paper are not publicly available at this time but could be obtained from the authors upon reasonable request.

Supplementary information

Supplementary information for this paper is available at https://doi.org/10.29026/oea.2026.250312