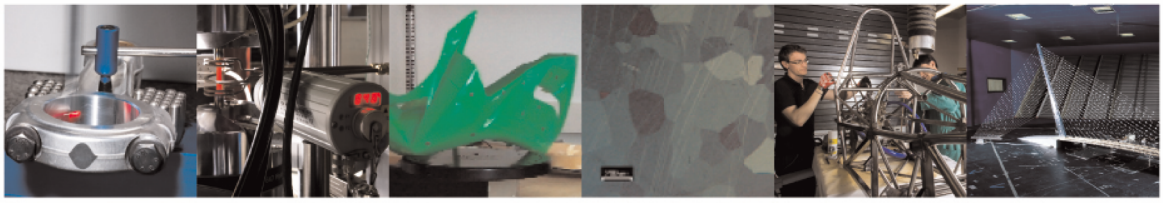

# The synergy of synchrotron imaging and convolutional neural networks towards the detection of human micro-scale bone architecture and damage

F. Buccino;I. Aiazzi;A. Casto;B. Liu;M. C. Sbarra;G. Ziarelli;G. Banfi;L. M. Vergani

This is a post-peer-review, pre-copyedit version of an article published in Journal of the Mechanical Behavior of Biomedical Materials. The final authenticated version is available online at: https://doi.org/10.1016/j.jmbbm.2022.105576



DIPARTIMENTO DI MECCANICA ■ POLITECNICO DI MILANO
via G. La Masa, 1 ■ 20156 Milano ■ EMAIL (PEC): pecmecc@cert.polimi.it
http://www.mecc.polimi.it
Rev. 0

# The synergy of synchrotron imaging and convolutional neural networks towards the detection of human micro-scale bone architecture and damage

Federica Buccino[a], Irene Aiazzi[a], Alessandro Casto[a], Bingqi Liu[a], Maria Chiara Sbarra[a], Giovanni Ziarelli[b], Giuseppe Banfi[c], Laura Maria Vergani[a]

[a] Department of Mechanical Engineering, Politecnico di Milano, Italy

[b] Department of Mathematical Engineering, Politecnico di Milano, Italy

[c] IRCCS Istituto Ortopedico Galeazzi, Milano, Italy

* Corresponding author. Tel: 02-2399.8249 Email: laura.vergani@polimi.it (Laura Maria Vergani, Via La Masa 1, 20156 Milano, Italy)

## Abstract

The growing health and economic burden of bone fractures, their intricate multiscale features and the existing knowledge gaps in the comprehension of micro-scale bone damage occurrence make fracture diagnosis a challenging issue. In this scenario, deep-learning and artificial intelligence embody the new frontier of healthcare system, by overcoming the subjectivity of clinicians in the analysis of medical images. However, the preliminary attempts in exploiting the power of machine learning algorithms such as neural networks are still limited to bone macro-scale, while there is an evident lack in their application to smaller scales, where damage starts nucleating. Currently, speculations at the micro-scale are only feasible with the aid of high-resolution imaging techniques, that are particularly time consuming in terms of output images analysis. In this context, this works aims at combining the visualization of the micro-crack propagation mechanism with the promising application of convolutional neural networks. The implemented artificial intelligence tool is based for the first time on a large number of human synchrotron images coming from healthy and osteoporotic femoral heads tested under micro-compression. The designed convolutional neural networks are able to automatically detect lacunae and micro-cracks at different compression levels with high accuracy levels; indeed, with the baseline setup, networks achieve more than 0.99 level of accuracy for both cracks and lacunae, and more than 0.87 of the meanIoU adopted as validation metric. This approach is particularly encouraging for the development of powerful recognition system to comprehend bone micro-damage initiation and propagation, paving the way to the application of machine learning studies to bone micromechanics. This could be additionally crucial for future patient specific micro-scale observations to be related to the clinical practice.

## Keywords

Lacunae; micro-damage; osteoporosis; convolutional neural networks; synchrotron; accuracy

## 1. Introduction

Bone fractures are becoming a growing health and economic worldwide burden. As an example, in the United States, annual fractures are expected to rise by 50% from 2005 to 2025: this results in a consequent dramatic increase of the associated costs in the considered time interval [1]. This significant issue is closely linked to the population's average age increment. Indeed, with maturity, human bone appears more fragile and easily susceptible to fracture. The leading bone fracture causes could be traumatic events or pathologies, among which the most common is osteoporosis which has a massive incidence on the population, especially in women, resulting in a dramatic reduction of bone mineral density. As a consequence, bone fractures and osteoporosis prevention, early diagnosis and treatment, are crucial in order to reduce the impact of these pathologies on health care system.

The comprehension of fracture mechanisms is very challenging due to the complexity of bone structure [2]. Indeed, bone consists of a deeply sophisticated hierarchical architecture. At the macro-scale (50 cm-10 mm) it is possible to distinguish the whole bone structure, while the meso-scale (10 mm-500 μm) allows to appreciate two different kinds of bone: cortical and trabecular arrangements [3,4]. The microscopic level of the bone (500-1 μm) is composed by different arrangements of lamellae. They are the result of mineralized collagen fibrils assembled in a sheet-like structure approximatively 3-7 μm thick [5]. In the cortical bone tissue, lamellae are arranged into concentric layers forming osteons (Haversian systems). Along osteon axis, a microscopic tube with a diameter around 200 μm [6], named Haversian canal, contains blood vessels and nerves. Neighbouring Haversian canals are interconnected through vascular transversal channels, i.e. Volkmann's canals, having the function of nourishing and interconnecting cells [7]. At this level, it is also possible to distinguish lacunae, biconvex lens-shaped cavities where osteocytes lie. Instead, concerning the trabecular bone tissue, trabeculae are formed by a web of aggregated lamellar bone, full of ellipsoidal micro-porosities, i.e. lacunae. Lacunae are connected by dendritic channels called canaliculi, which compose the lacuno-canalicular network, allowing nutrients and waste to flow from one osteocyte to the next [8]. Specifically, their contribution in both strength and toughness is still debated. More in depth, there exists a dualism behind the role of lacunae: they seem to play an antithetical mechanical contribution, having an effect on both strength and toughness [9–13]. In the first instance, lacunae could be classified as stress concentrators [14–17], by representing bone discontinuities able to locally amplify stresses and strains. In this sense, lacunar system should contribute to strength reduction. However, considering bone as a damage-tolerant material [18], lacunae are able to deviate the crack path, slowing down damage progression [19,20] and consequently increasing material toughness. At the nanoscale (50-0.5 nm) it is possible to observe the main components of the bone such as collagen fibrils, hydroxyapatite crystals, non-collagenous proteins, proteoglycans and water [21]. Fracture patterns reflect this complex architecture and damage processes affect all the different hierarchical levels [22].

This entangled structure makes also fracture detection mechanisms particularly intricate. While there are several studies regarding damage progression at the macro and meso-scale [5,23], it is still unclear what occurs at the micro-scale and what is the role of lacunae during damage processes. Currently, medical images assess to be the main instrument to judge bone fracture occurrence [24]. With the advancement of medical technologies, there are several options for obtaining high-quality medical pictures, including X-rays, computer tomography, magnetic resonance imaging, and ultrasound. At the macro-scale, it is a common practice to visually examine an X-ray picture to assess the existence and severity of a bone fracture in order to administer a suitable treatment. Nevertheless, limitations arise in the identification of position and orientation of a bone fracture with traditional X-ray imaging. This process requires an accurate training of the medical staff, it is strongly subjective, it is often affected by human errors and it is particularly time consuming [24].

In this scenario, computer assisted diagnosis represents a promising tool in the analysis of medical images to assist clinicians in the detection of bone fractures, and it has garnered a growing amount

of attention during the last years [24]. In fact, digital medicine, deep-learning and artificial intelligence embody the new frontier of healthcare system. More in depth, artificial Neural Networks (ANN) start to be used to improve diagnostic process and to reduce the risk of misdiagnosis [25].

There are preliminary existing attempts in exploiting the power of neural networks in macro-scale bone fracture, especially in the context of osteoporotic disease. As an initial approach, Pradhan et al. [26] have implemented a deep convolutional neural network (CNN) to elaborate a new classification method of human bones based on X-rays image analysis. The encouraging results that emerge from their analysis pave the way for further investigations as a support for the diagnosis. As an example, given the lack of an absolute method for diagnosing osteoporosis, plenty of works in literature propose neural networks capable of detecting this pathology based on image analysis and combined with the indices and criteria defined by the World Health Organization for osteoporosis and fracture risk, such as T-score examination [25]. Liu et al. [27] have conceived an ensemble ANN to point out the most important risk factors for hip fractures, obtaining similar results with respect to other related works, thus they have demonstrated that ANN assesses to be a remarkably efficient prediction instrument [27]. Yang et al. [28] have succeeded in realising a contour-based fracture detection scheme with a good accuracy (74.4%) and sensitivity (82.9%), that is able to identify fracture lines, their coordinates, midpoints and slopes. CNNs are also used for image segmentation and classification of osteoporotic fractures in vertebral bones [29] while similar methods are used to assess bone age [30]. A neural network has also been implemented to segmentate automatically ultrasound images [31] using a Fully CNN on a sample of bones from different volunteers. Similarly, image recognition through CNNs is performed on Computer Tomography (CT) [32] and Magnetic Resonance (MR) [33] images. A comparison among several ANN approaches applied to macro-scale bone images is reported in Table 1.

Table 1: Bone CNN recognition with a comparison among the aims of the ANN application and the method of evaluation for the network. Accuracy is calculated as the number of correct predictions over the total number of predictions. Dice Score is a metric that indicates the validity of the predicter output with respect to a reference (ground) and it is calculated as: $\text{Dice Score} = \frac{2\,|O_p \cap G_t|}{|O_p|+|G_t|}$ with $O_p$ predicted output and $G_t$ ground truth.

| Authors | Method | Aim | Evaluation |
|---|---|---|---|
| Rehman F. et al. [29] | U-Net, Python with Tensorflow | Segmentation of vertebra images | Dice Score = 0.96 ± 0.01 |
| Villa M. et al. [31] | Fully CNN, Python in the Caffe environment | Recognition of various types of bones (femur, tibia and pelvis) | Accuracy = 80% |
| Hemke R. et al. [32] | CNN, Python, Keras library and Tensorflow | Segmentation of axial images of the pelvis | Dice Score = 0.92 |
| Pradhan N. et al. [26] | Deep CNN Network | Recognition of different types of bones | Accuracy = 91.37% |
| Liu Q. et al. [27] | Three-layer back-propagation ANN | Prediction of the role of different factors leading to hip fracture taking as training dataset surveys on patients. | Accuracy = 85% for females; 93% for males. |
| Yang A. Y. et al. [28] | ANN applying Principal Component Analysis | Detection of long-bone fracture based on line features of X-ray images | Accuracy = 74.4% |
| Zengya B. et al. [30] | Deep CNN with Tensorflow | Detection accurately bone age | Accuracy = 91.3% |
| Deniz C. M. et al. [33] | Two different CNN | Automatic segmentation of Femur from MR images | Dice Score = 0.95 ± 0.02 |
| Chin-yun Shen et al. [34] | CNN, residual neural networks and transfer learning | Classification and prediction of mechanical states of cortical and trabecular bone tissue | Accuracy = 98% for cortical bone; 90% for trabecular bone. |

Chin-yun Shen et al. [34] are the first to recently apply machine learning methods to analyse the micro-scale architecture by means of high-resolution synchrotron-radiation micro-computed tomography. They employ CNNs, residual neural networks, and transfer learning to classify and predict mechanical states of bovine cortical and trabecular bone tissue investigating micro-computed tomography images acquired in uniaxial continuous compression. They develop trained models that identify new images between failed and pristine classes with over 98% accuracy for cortical bone and over 90% accuracy for trabecular bone using convolutional neural nets architectures. This shows that even with a small number of training samples (5 samples), powerful classifiers for high-resolution micro-CT images can be developed, and it encourages further development by including more data and training methods to deepen machine learning-driven insights into microstructural states and properties of bone.

The increasing interest in applying ANN for the study of multi-scale bone damage lies in the possibility to assist clinical doctors in their diagnosis. In particular, the work by Lindsey et al. [35] has demonstrated that when emergency medicine doctors are supported by a deep learning trained model, their ability to accurately detect fractures improves significantly: the average clinician's sensitivity is 80.8% without the help of artificial intelligence and 91.5% when they are aided by the model. Through this kind of tool, it is also possible to assess the effect of a therapy, as demonstrated by Messina et al. [36]. Additionally, ANNs are evolving very rapidly in the healthcare world as they are proving to be a highly suited mean to perform more efficient and less time consuming image segmentation and classification [36].

Nevertheless, no one has yet exploited the power of this artificial intelligence tool applied to the detection of human micro-scale bone damage. This is mainly because the contribution of micro-scale architecture to micro-damage is still debated, even if several authors have addressed the discussion by exploiting experimental [16,37,38] and numerical approaches [39–41]. If the microcracks are not so severe, bone can withstand plenty of them, behaving as a damage-tolerant material [42]. Microcracks generally take origin in high stress region and micro-scale heterogeneities impact the crack's propagation. In this perspective, the lacunar network, and its antithetical effect on both strength and toughness, has recently received increasingly more attention [43].

Histological sections, which lack three-dimensionality, have traditionally been used to examine micro-architectural properties, with the aim to find correlations with macro-scale cracks development. With the advent of much more sophisticated micro-CT image technologies, which have a resolution capable of capturing not just trabecular architecture, but also ultrastructural porosities, including lacunae, the aforementioned phenomena could be highlighted and possibly clarified. However, dealing with this type of high-resolution images results in very high computational costs linked to both the acquisition, their processing and segmentation, filtering and reconstruction steps, that are significantly time consuming. Moreover, other typical issues related with the segmentation of micro-scale bone images come from the spatial complexity of the structure and the variability of microcracks. All the current criticalities encountered in the high-resolution images post-processing steps [44–51] are summarised in Figure 1.

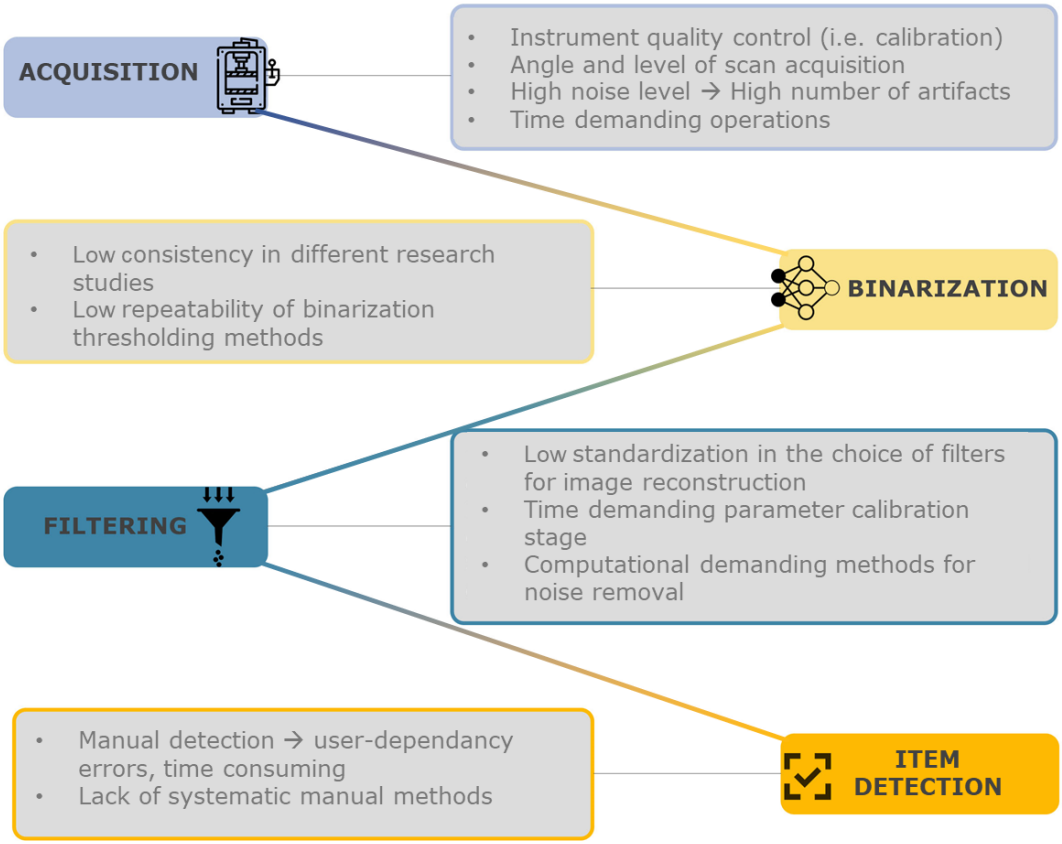


Figure 1: Schematization of the issues related to high-resolution images post-processing steps.

The current work exactly fits into this scenario and aims at combining the understanding of the micro-crack propagation mechanism and the promising application of CNNs. In this paper a novel implementation of a CNN to separately recognize both lacunae and micro-cracks of healthy and pathological human bone tissue images obtained with synchrotron light (X-ray) spectroscopy is proposed. One of the goals is to overcome some of the issues encountered in the processing stage, such as time-consumption and lack in the standardization of the choice of filtering, providing an automatic tool based on supervised learning. The processed high-resolution outputs can lead to increase the understanding of the role of micro-cracks and their specific interaction with bone micro-structure. The novelty of this work is that the development of the artificial intelligence tool is based on a large number of human high-resolution synchrotron micro-CT images in order to reach a better accuracy with respect to previous literature works [34]. It has a great potential since similar applications at larger scales, as demonstrated in Table 1, have shown that CNNs are very suitable to deal with classification and detection issues.

## 2. Materials and Methods

Trabecular samples obtained from healthy and osteoporotic human femoral heads (FH) are considered in this study and tested with an ad hoc micro-compression device inside a synchrotron for the evaluation of micro-damage initiation and progression. The output images from the synchrotron are consequently processed by means of CNNs. The schematic of the whole methodological approach is reported in Figure 2.

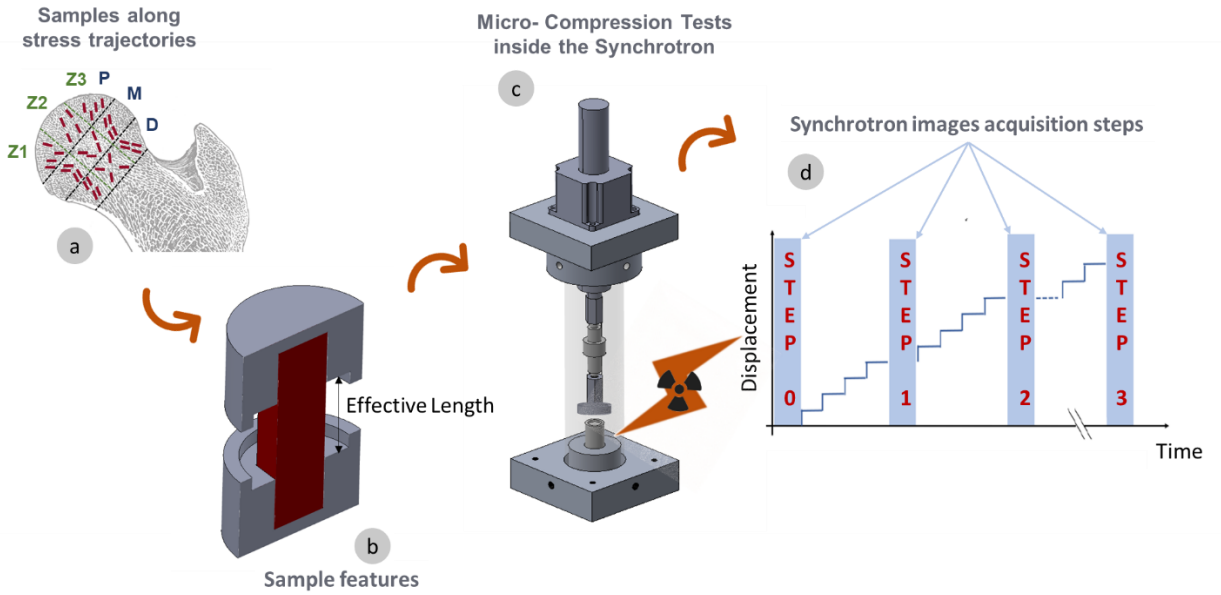


Figure 2: methodological approach. (a) Samples from healthy and osteoporotic patients are cut along the principal stress trajectories. (b) samples are embedded into epoxy resin endcaps, defining sample effective length as the distance between the upper and lower endcap. (c) Samples are then tested with a micro-compression device under displacement control inside a synchrotron. (d) multiple synchrotron images are acquired at different displacement levels.

## 2.1 Samples preparation

Femoral heads are collected prior authorization from the Ethics Committee (approval date: 13/05/2020, ClinicalTrials.gov ID: NCT04787679) of San Raffaele Hospital (Milano, Italy) and signed approval consent of the patients. The patients are over eighteen years old and are eligible to femoral neck osteotomy for primary hip replacement. Samples are obtained by cutting the femoral head along the principal tensile and compressive stress trajectories (Figure 2a) and are classified with respect to their positioning, in accordance to a specific partition in zones (Z1-in the frontal plane, Z2 and Z3-in the posterior plane) and sub-zones (Proximal, P, Medial, M, and Distal D with respect to the pelvis positioning). The whole partitioning process in nine sub-regions is reported in detail in a previous work [52]. The specific characteristics of each sample considered in the present work are reported in Table 2.

Table 2: specific characteristics of each sample: positioning, patient age, radiologic information, and additional clinical observations. The definition of osteoporosis reported in the column “Radiologic information” is in accordance with the World Health Organization (WHO) guidelines.

| Sample ID | Sample positioning | Patient age [years] | Radiologic information | Additional observations |
|---|---|---|---|---|
| FHC1_S30 | Z1_M | 62 | Non osteoporotic | No other comorbidities |
| FHC1_S57 | Z2_P | 62 | Non osteoporotic | No other comorbidities |
| FHC1_S59 | Z2_D | 62 | Non osteoporotic | No other comorbidities |
| FH5_S33 | Z2_M | 76 | Non osteoporotic | Osteopenic |
| FH1_S7 | Z1_M | 61 | Non osteoporotic | Large bone spurs |
| FH4_S31 | Z2_M | 84 | Osteoporotic | / |

Samples are then embedded into epoxy resin endcaps: the distance between the upper and the bottom endcap is defined as effective length (L_eff) of the sample, corresponding to the unconstrained bone portion (Figure 2b). Samples, that have a square cross-section of 16 $mm^2$ and a total length (L_tot) of 14 mm, are fixed in formaldehyde to avoid bacteria contamination and then

stored in 70% ethanol. From each femoral head, a total of 4 samples is obtained from each of the nine sub-partitions. The samples that are not mentioned in Table 2 are tested under compression outside the synchrotron, to allow for a global mapping of the FH mechanical characteristics.

## 2.2 Mechanical tests

The testing phase conducted inside the synchrotron requires the design of an ad hoc Micro-Compression Device (MCD) (Figure 2c). The technical specifications of the MCD are reported in [52]. In particular, the device is characterized by limited weight (<5kg) and height (<50 cm), stability, ease of disassembling and X-ray transparency.

Before the mechanical compressive tests, performed under displacement control, the trabecular samples are rehydrated in saline solution for one hour, then are kept wet during the test to simulate the in-vivo environment. Bone samples are, then, pre-loaded (2-3 N) with 3 consecutive cycles raising and dropping the motor plunger by 0.1 mm from the zero-load position. Each compression step is performed with a step amplitude of 0.1 mm, a displacement rate of 0.1 mm/s and is followed by 3 minutes of relaxation time in which the imposed displacement is maintained. The instantaneous load measured by the cell is recorded with a sampling frequency of 10 Hz and plotted in real-time allowing it to be continuously monitored during the test. A specific alternance of tomographic acquisitions and micro-compression steps is proposed for the tested specimens, according to Figure 2d. In addition to the micro-compression test, which takes on average 40 minutes, each tomographic acquisition lasts around 50 minutes, for a total of about 4 hours per sample. Mechanical curves of the six considered specimens are reported in Figure 3a, while Figure 3b shows values of the Young modulus of the four FH, normalized by the maximum value detected in each FH. This normalization permits to neglect intrinsic variations linked to the donor age range or additional alterations in the FH morphology that are not dependent to the physio-pathological condition. In addition, the effect of synchrotron radiation exposure on the mechanical characteristics of the samples is checked: the mechanical properties of the six scanned samples are within the values detected in the respective femoral head during mechanical tests outside the synchrotron (Figure 3b); thus, validating further considerations on scanned samples. There's not a unique and homogeneous trend evidencing the change of properties due to sample irradiation, leading to the conclusion that the obtained values are linked predominantly to intra-patient variability. Concerning post-yield properties, statistically significant differences ($p<0.05$) could be detected in terms of normalized yield strength and ultimate strain for the osteoporotic FH4. The normalized ultimate strength shows differences for osteoporotic samples belonging to FH4 and for non-osteoporotic FH1 and FHC1. Indeed, the irradiation affects specifically the post-yield mechanical behavior [53], resulting in a progressive loss in the post-yield deformation and a decline in strength, toughness, and ductility of bone.

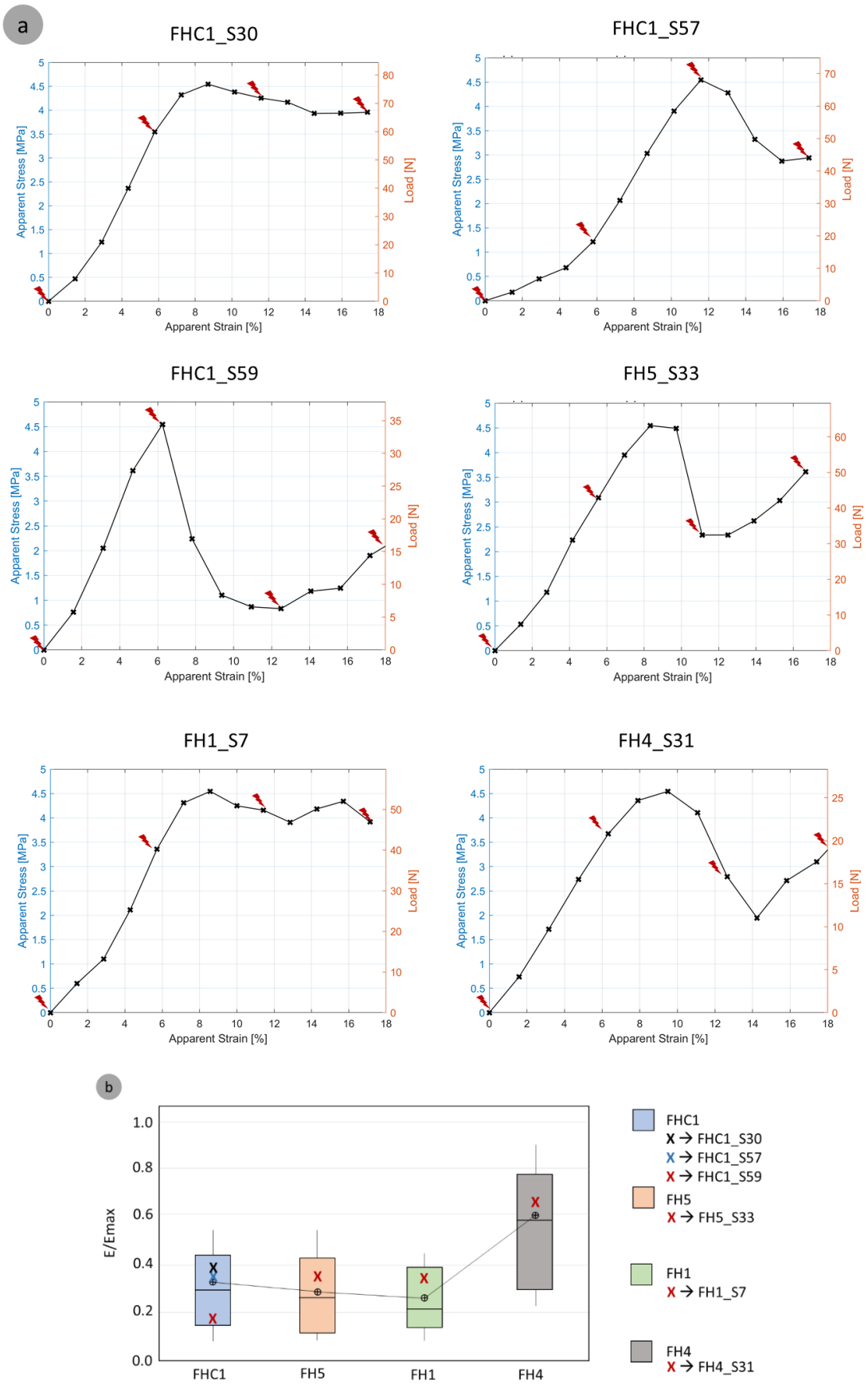


Figure 3: mechanical characteristics of the samples. (a) mechanical curves in terms of apparent stress-apparent strain-load of the six scanned samples. Red bolts represent synchrotron scans. (b) normalized Young's modulus for the FH to which scanned samples belong. Crosses represent the scanned samples.

## 2.3 Synchrotron imaging

The micro-compression tests are performed in the experimental hutch, placed inside ELETTRA Synchrotron (Trieste, Italy). The synchrotron monochromatic beam (electron beam: 2 GeV, 158 mA) is emitted (Figure 4, element a) at the energy of 25.6 keV after being filtered by the presence of a silicon foil (1.5 mm thick) which eliminates the low energy spectrum components, increase the signal to noise ratio, and crosses the human trabecular bone specimen, immersed in saline solution inside the inner cylinder of the MCD (Figure 4, element b). This latter is rigidly coupled through four screws to the underneath rotatory platform (Figure 4, element c), which can perform rotations of 360°. The scintillator, installed behind the MCD (Figure 4, element d) at a distance of 15 cm, acquires the high-resolution images. The linear actuator is connected to an Arduino UNO board and remotely commanded from the control room during the mechanical test.

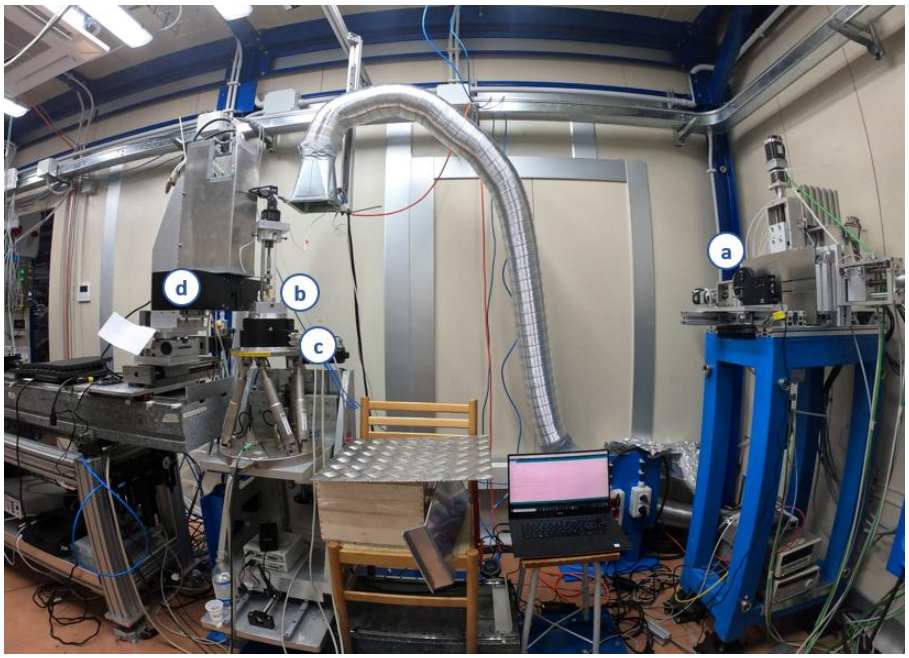


Figure 4: Experimental setup at ELETTRA Synchrotron. The synchrotron light is emitted (a) and hits the bone specimens inside the MCD (b) which is placed on the rotatory platform (c). Tomographic images are acquired through the scintillator installed behind the MCD at the sample level (d).

The tomographic acquisitions are performed according to the parameters detailed in Table 3.

Table 3: chosen parameters for synchrotron images acquisition.

| Acquisition parameters | Acquisition parameter's value |
|---|---|
| Beam energy [keV] | 25.6 |
| Si thickness [mm] | 1.5 |
| Scintillator – sample distance [mm] | 150 |
| Pixel size [µm] | 1.6 |
| Field of view [mm] | 3.28 · 3.28 |
| Range of scanning [°] | 360.2 |
| Projections | 1800 |

The major consideration in setting these parameters is the trade-off between scanning time and image quality: faster scans reduce the delivered radiation dose and sample motion issues, but also have lower signal-to-noise ratio due to few X-rays reaching the detector for each image. The scan is performed in fly mode, indeed the MCD continuously rotates on the platform during the acquisition and one flat field image (without the sample) ad one dark field image (without X-

ray illumination) is taken. Given that the dimension of the electron beam is not enough to cover the entire L_eff of the sample, 3 consecutive sub-scans with a step of 3 mm are acquired.

Tomographic projections are acquired and stored in the raw file .TDF during the experimental campaign and then properly processed. The first stage is the reconstruction, in which the raw files are converted into a set of subsequent bidimensional images of the volume. Reconstruction is a high computationally and memory-demanding phase: the raw .TDF file associated with a sub-set has a dimension of about 26 GB. This file must be converted into a set of 2048 squared images with dimensions in the range of 3500-4000 pixels for each side, leading to about 150 GB dataset. Moreover, a single specimen requires 3 tomographic acquisition for each of the 4 compression steps (Figure 2d). When processing SRμCT data, artefacts from the scanning process must be removed during the reconstruction to ensure an accurate material characterization. The main encountered scanning artefacts include stripe or ring artefacts, beam hardening, metal or attenuation coefficient artefacts, and shifted rotation center. Stripes in the original scanned data appear as rings in the reconstructions and are a result of inhomogeneous X-ray beam drift. Phase retrieval techniques can be used to quantitatively reconstruct pixel grey values corresponding to the phase contrast, by solving the transport of intensity equation, as reported in [54]. The reconstruction is performed through SYRMEP Tomo Project (STP) v 1.4 [55], an open-source software used and implemented at the SYRMEP beamline. The considered reconstruction parameters chosen for sample processing are reported in Table 4. The sample rotation axis is parallel to the direction of the applied displacement, and the attenuation of the X-ray beam occurs in the perpendicular direction. Therefore, the reconstruction via STP software is by default perpendicular to the rotation axis, that implies the minimum level of complexity [56]. Reconstruction along other directions are more computationally demanding and could be eventually obtained in a post-processing phase.

Table 4: Reconstruction's parameters chosen for samples processing stage.

| Sample ID | Extended Field of View Overlap [pixels] | Phase retrieval | Angle projections | Ring Removal Parameter |
|---|---|---|---|---|
| FHC1_S30 | 529 | 20 | 180 | 11 |
| FHC1_S57 | 410 | 20 | 180 | 11 |
| FHC1_S59 | 437 | 20 | 180 | 11 |
| FH5_S33 | 906 | 20 | 180 | 11 |
| FH1_S7 | 363 | 15 | 180 | 11 |
| FH4_S31 | 552 | 18 | 180 | 11 |

After the reconstruction, some noise could remain present in the inter-trabecular voids. Additionally, the manual detection of lacunae and cracks is particularly challenging and time-consuming, and the computational costs are significantly high.

## 2.4 Manual labelling process

Supervised learning algorithms, such as the implemented neural network, need a sufficient amount of manual pre-labelled images. Original reconstructed images are saved in .TIFF format, 32-bit single channel. In order to make this step automated and valid for all the acquired scans, a Matlab code is implemented which apply the following operations to 2D slices: 1) Gaussian filter with 2.5 standard deviation; 2) Binarization, generating a mask binary image: pixels over a threshold are considered full, otherwise empty. The threshold choice is based on the Otsu algorithm implemented in the Matlab function imbinarize; 3) Opening operation, using a square-shaped kernel with a dimension of 15 pixels; 4) Closing operation, using a disk-shaped kernel with a dimension of 20 pixels. The higher dimension and different shape are chosen to smooth contours and prevent the presence of voids

after this step, finally obtaining a mask; 5) Mask application: the mask obtained in the previous step is applied directly to the original images. Voids are taken as 0-valued pixels, while inside the bone region, the exact value of the original image pixels is maintained. After the filtering stage, cracks are manually detected in each slice, and the image is converted in Portable Network Graphics format. The manual labelling process is performed by six different operators; the average operator error is about of 1/100 recognized cracks, estimated from a cross-validation between the operator detection process. As concerns segmentation of lacunar structures, since the manual recognition is extremely time demanding, due to the high number of lacunae recognizable in each slice, an algorithm is implemented and reported in [57]. All objects smaller than 50 $\mu m^3$ and larger than 2000 $\mu m^3$ are removed as to reflect the range of human lacunar volumes reported in previous studies [58–61]. Eventual artifacts escaping the volumetric filter, and all exhibiting similar thin configurations with a high object-elongation value are excluded too (elongation>0.85). This algorithm is proven to be accurate (manual identification is used as gold-standard for calculating accuracy) for lacunar identification, reproducible for repeated measures, and sensitive for lacunar parameters between anatomically distinct regions [57].

## 2.5 Neural networks' implementation

A Convolutional Neural Network (CNN) is implemented for cracks and lacunae detection, based on the Keras VGG16 built-in model [62]. This model has achieved 92.7% accuracy of top-5 test on the ImageNet dataset, which is a classification problem based on image recognition of over 14 million images belonging to 1000 distinct classes. It has been built to improve the previous existing AlexNet model by replacing kernel-sized filters with main working blocks as multiple 3x3 kernel-sized convolutional filters. 16 stands for the total number of convolutional layers which constitute the network. In classification problems where the output must correlate images to distinct classes, it is suggested to adopt a SoftMax activation function in order to convert numerical outputs in probabilities of belonging to a specific class.

Our scenario adopts the VGG16 model in order to prescribe each image with a relative mask of the same dimension, highlighting some sites of interest such as cracks and lacunae. In Figure 5 a schematic representation of the topology of the CNN is shown. The output of the implemented architecture is a matrix which dimensions are the same of the input image, where each entry is a value ranging from 0 to 1 after the softmax layer has been applied. Therefore, applying the last filter layer a black/white mask is obtained, where white pixels identify cracks or lacunae respectively.

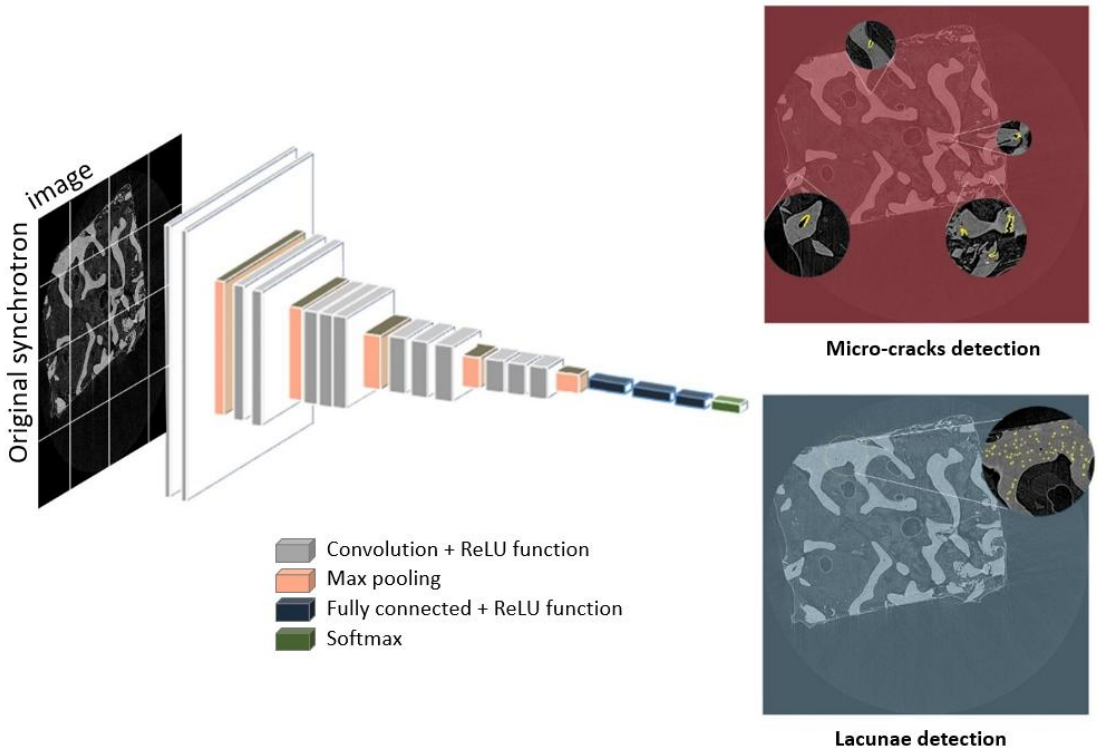


Figure 5: Schematic representation of the two CNNs for detecting cracks and lacunae. The colored parallelepipeds represent the building blocks of the network. Grey stands for the application of the convolution combined with Rectified Linear Unit (ReLU) function, pink for max pooling, blue for fully connected combined with ReLU function and green for softmax application.

Images reconstructed from Elettra Synchrotron are reduced to 1858 X 1858 pixels due to limitations for image acquisition and recognition. These images are tiled into 16 parts in order to ease the training stage of the Neural Network. Each of the 16 tiled images has four color channels (RGB with transparency) since the original one is a PNG file, and the dimension is 464 X 464 pixels, corresponding to the set of input images to the convolution layer that is used to feed the Network. Images, along with their corresponding mask labels, are then shuffled to avoid batches full of background. Additionally, the ones corresponding to black masks are then deleted in order to guarantee a faster convergence of the network. The output of the network is a mask of the original image with dimension 464 X 464 pixels. We also considered a 9-tiling procedure of the same images for both training and testing stage. However, dealing with smaller dimension images has outcome in better accuracy of the Network, as highlighted in paragraph 3. For the prediction stage, which will require the assembling stage of the 16 tiled images, there is the need of an untiling function to regather the sub-images. Adam optimizer is implemented in Keras, which is a stochastic gradient descent approach based on adaptive estimation of first-order and second-order moments. The method is particularly suitable for the purpose of this work, being it computationally efficient, with little memory requirement and suitable for problems that are large in terms of data/parameters [63]. As for the hyperparameters to be tuned, we initially fix the learning rate at 1e-4 as a baseline. The validation metric that has been chosen is the mean Intersection over Union (meanIoU), which consists in the evaluation of the IoU for each semantic class (in our case black and white) and then computing the average over classes. IoU is defined in Equation 1.

$$IoU = \frac{True\_Positive\_pixels}{True\_Positive\_pixels + False\_Positive\_pixels + False\_Negative\_pixels} \quad \text{Eq.1}$$

Some hyperparameters cannot be inferred while fitting the network to the training set and have to be tuned properly in order to improve performances of the CNN in terms of accuracy and meanIoU. In this work different simulations have been run so as to deduce suitable values for the learning rate, that is the rate at which weights and biases of the network change, and the minibatch size, which is the number of samples whose gradients is computed during each epoch iteration.

### 2.6 Neural network training

Two different training stages have been addressed in this work, one for crack and one for lacunae detection.

As regards micro-cracks recognition, during training and validation 648 images are used (randomly split between validation: 20% and training: 80%). Indeed, the input set is constituted by six different samples from four different individuals, i.e. we dealt with 108 different slices belonging to FH1, FH4, FH5 and FHC1 (three samples come from the same FHC1). The specific choice of considering four different individuals reside in the need to avoid overfitting of the model. We selected different slices from those individuals in order to reduce the generalization error of the network across different specimens and different compression steps (see Table 5). Except for FH4, the other samples belong to healthy individuals. The imbalance between healthy and osteoporotic subjects in the choice of training images is geared towards the purpose of the research, namely, to provide recognition of cracks in the early stages of fracture, when cracks are difficult to be detected.

Concerning lacunae identification, the input set is constituted by the same dataset described for crack recognition. Samples are added one by one in order to reach more than 0.87 of meanIoU.

Table 5: Samples for the training stage for micro-cracks and lacunae detection.

| Training for the detection of: | Sample ID | Number of slices | Compression steps |
|---|---|---|---|
| Micro-cracks and lacunae | FHC1_S30 | 108 | 1,2,3 |
| | FHC1_S57 | 108 | 1,2,3 |
| | FHC1_S59 | 108 | 1,2,3 |
| | FH5_S33 | 108 | 1,2,3 |
| | FH1_S7 | 108 | 1,2,3 |
| | FH4_S31 | 108 | 1,2,3 |

## 3. Results

### 3.1 Convergence analysis

In the whole proposed analysis, the topology of the VGG16 is not altered: the neural network is constituted by 2-2-3-3-3 convolutional layers alternated by max-pooling layers. For the choice of the hyperparameters, several simulations have been run with different parameter settings, taking under control both accuracy of the Network and meanIoU for the validation set. Convergence is explored during the training by tuning two different parameters as mentioned in Section 2.4: the learning rate and the minibatch size for the Adam Optimizer. Convergence results in Figure 6a are related to the training stage performed for cracks detection, which showed more critical issues. First, we consider different learning rates, specifically 1e-1,1e-2, 1e-3, 1e-4, fixing the minibatch size of the implemented Adam Optimizer at 16.

It is notable that the maximum accuracy level is reached with learning rate fixed at 1e-2 (Accuracy= 0.9979) and the same trend is obtained with the mean-Intersection-over-Union (meanIoU=0.8743).

The very high level of reached accuracy is not associated with the risk of overfitting. Indeed, in the same cases meanIoU (which is greater than 0.5) achieves values that assure reliable predictions [64]. An increased accuracy level is localized in zones of no micro-cracks, characterized by black pixels. In order to delete such regions, corresponding to 0 gradient, that slow down convergence, we applied a tiling and untiling procedure. Hence, we split each 1858X1858 picture into 9 and 16 pictures and, in the training stage, we do not consider those sub-regions corresponding to a complete black mask. Since the 16-tiling configuration achieves better accuracy properties, from this point onward we perform simulations adopting this configuration (Table 6).

Table 6: Accuracy and meanIoU evaluated for two different tilings, i.e. dividing sample images into 9 or 16 sub-images.

| | **9-tiling** | **16-tiling** |
|---|---|---|
| **Accuracy** | 0.99756 | 0.99789 |
| **meanIoU** | 0.87441 | 0.87432 |

We compared performances of the network varying another hyperparameter, i.e. the minibatch-size. Setting learning rate to 1e-2, we analyse accuracy and meanIoU for three minibatch-sizes, namely bs = 8, 16, 32 (Figure 6b).

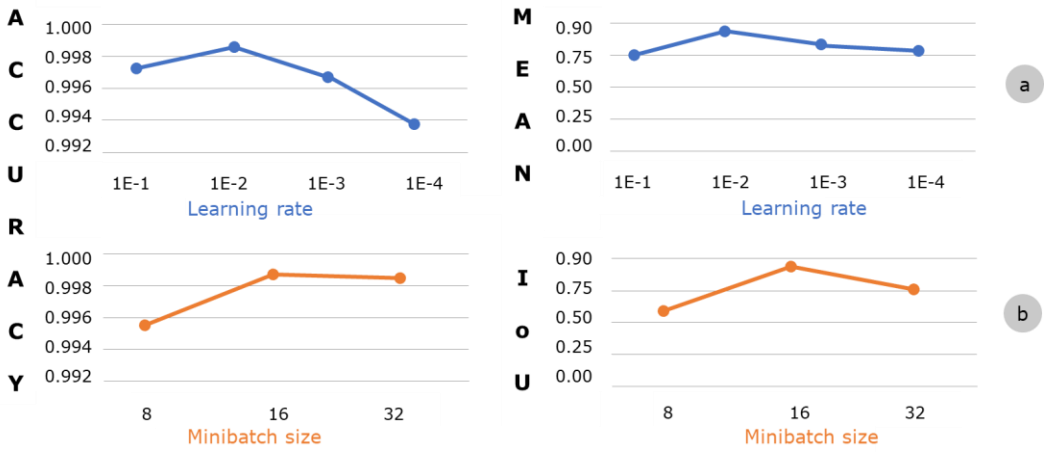


Figure 6: (a) Convergence results in terms of accuracy and meanIoU for CNN crack detection varying the learning rate. (b) Convergence results in terms of accuracy and meanIoU for CNN crack detection varying the minibatch size.

From these results, minibatch size is set to 16, in order to obtain higher reliability in terms of meanIoU which is a proper measure for VGG16 nets [65].

Analogous results are obtained when checking the convergence for the lacunae detection Neural Network (Figure 7 a and b): a learning rate equal to 1e-2 and a minibatch size of 16 return higher results in term of accuracy and meanIoU (0.8751).

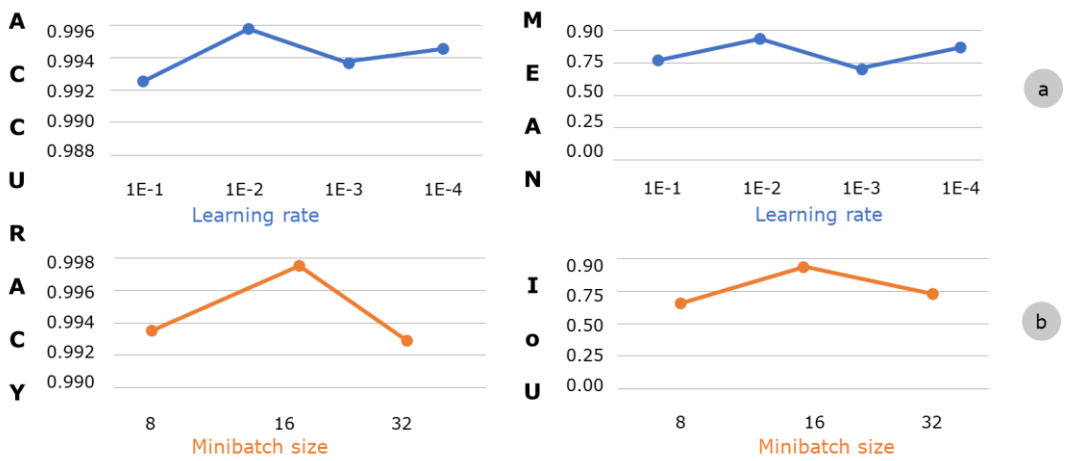


Figure 7: (a) Convergence results in terms of accuracy and meanIoU for CNN lacunar detection varying the learning rate. (b) Convergence results in terms of accuracy and meanIoU for CNN lacunar detection varying the minibatch size.

## 3.2 Micro-cracks detection through CNN

We tested the implemented Neural Network for detecting bone cracks in all the six samples scanned with synchrotron light at step1, step 2 and step 3.

Figure 8 presents an overview of different cases identified when applying the CNN for cracks detection. An accurate detection from CNN is present is Figure 8a, while Figure 8b and c shows possible manual detection errors induced by user. Manual detection errors could originate from several issues: tiny lateral cracks that are less visible in a manual detection process (as highlighted in Figure 8b) or an already compromised synchrotron image (as in Figure 8c).

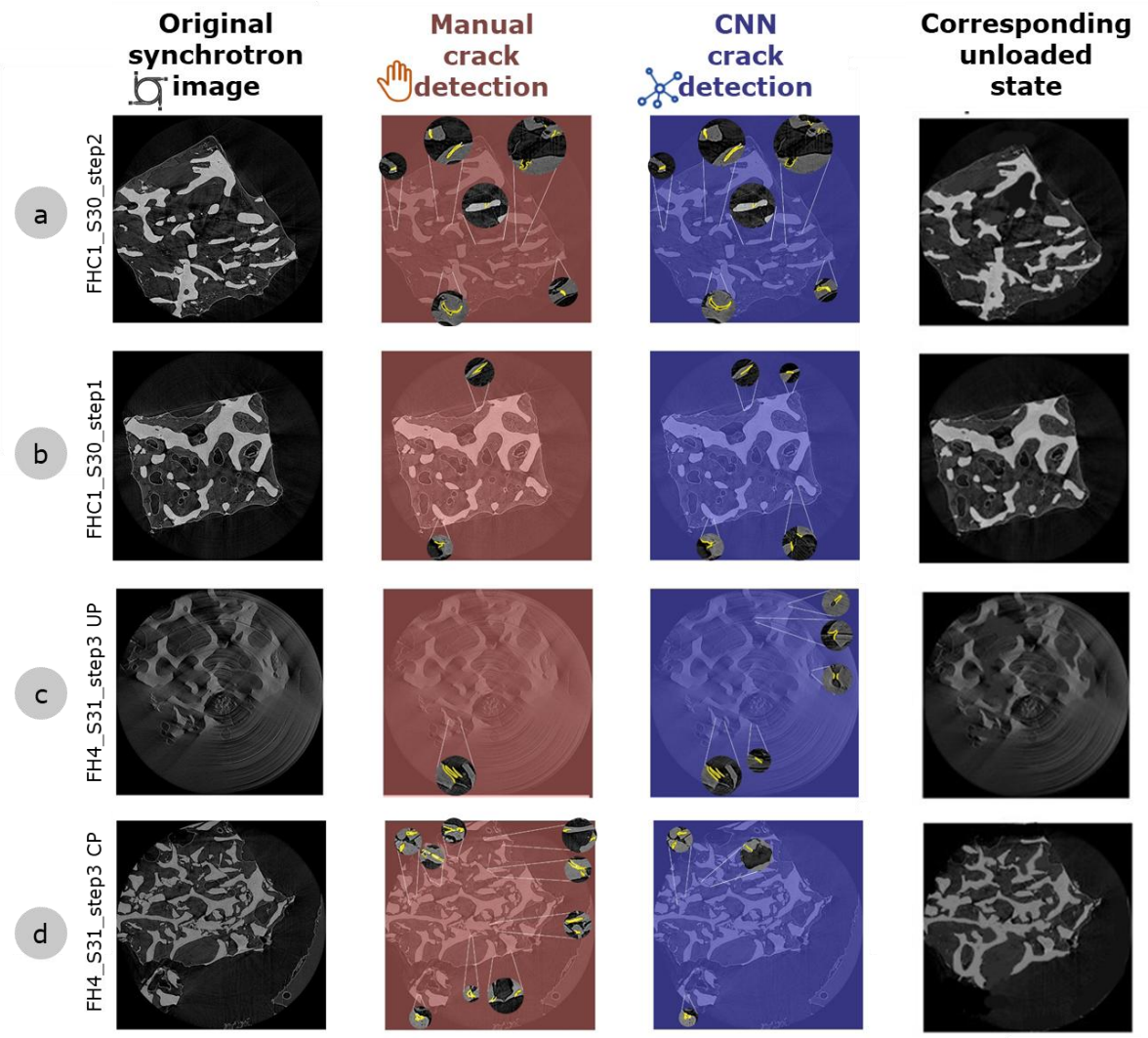


Figure 8: Micro-cracks detection (in yellow) by means of the implemented CNN (third column from the left). In the first column, it is reported the original image obtained by synchrotron's acquisition; in the second column the manual crack detection is presented. In the fourth column, the corresponding unloaded state is reported. (a) micro-cracks detection for the sample FHC1_S30 at the second compression step. (b) micro-cracks detection for the sample FHC1_S30 at the first compression step. (c) micro-cracks detection in the upper portion (UP) of sample FH4_S31 at the third compression step. (d) micro-cracks detection in the central portion (CP), closer to the loading plate, of sample FH4_S31 at the third compression step.

To obtain a quantitative measure of the network accuracy with respect to the user-detection, the error in crack identification evaluated at different compression steps for both manual and CNN approaches is presented in Figure 9.

We report in Figure 9 the micro-crack area evaluated at each compression step for each sample both by means of manual detection-based and AI-based approach. Only crack containing slices are considered in the evaluation. The relative error between manual and CNN-based segmentation is reported for each sample and for each compression level and it is defined as:

$$Relative\ difference\ \%_{Manual-CNN} = \frac{Average\ crack\ area_{Manual} - Average\ crack\ area_{CNN}}{Average\ crack\ area_{Manual}} . 100 \qquad \text{Eq.2}$$

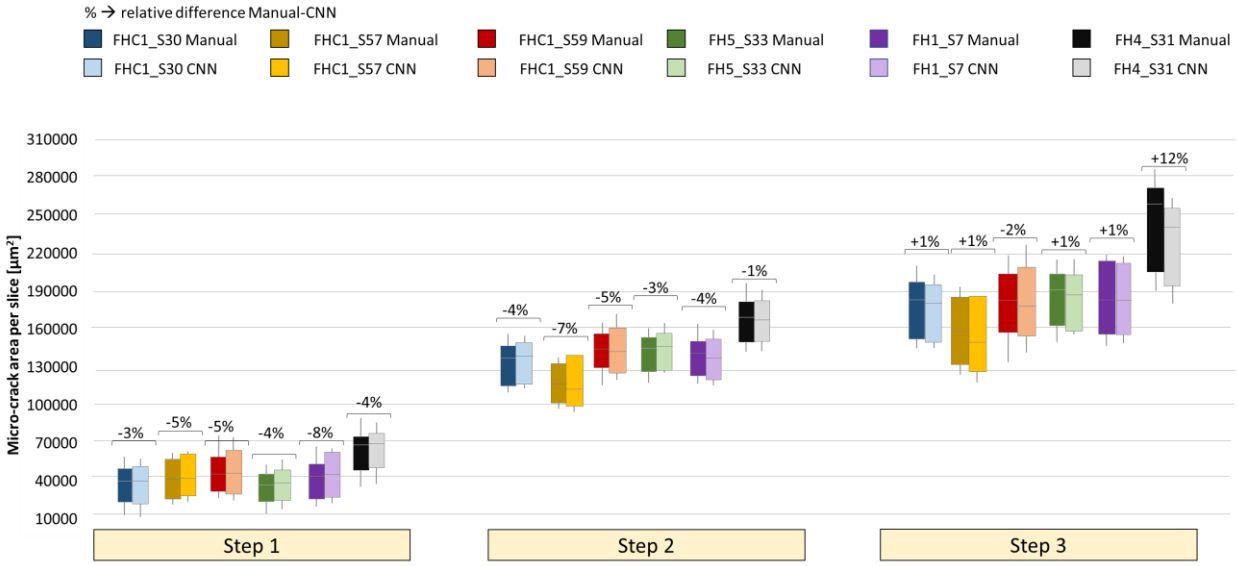


Figure 9: Values of micro-crack area measured both by manual and CNN-based detection. The error bars refer to the standard deviation of the analysed images for each sample at each compression level, considering the tiling process too. The percentage values are related to the percentage relative difference between manual and CNN-based crack recognition (Eq.2).

CNN errors in detecting micro-cracks increase in step 3 (Figure 8d), but it should be pointed out that the network succeeds in detecting bone cracks that were not identified by hand at initial compression steps (Figure 8 b).

### 3.3 Lacunae recognition through CNN

The CNN implemented for lacunae detection is tested on the same images and samples considered for crack recognition. Note that, with the same parameters chosen for micro-cracks detection CNN, the network succeeds in identifying a larger number of lacunae with respect to traditional approaches (Figure 10). Higher micro-scale porosity is detected in FH4 both by the CNN and the traditional detection strategy. The relative difference between manual segmentation and CNN-based segmentation is reported for each sample and for each compression level.

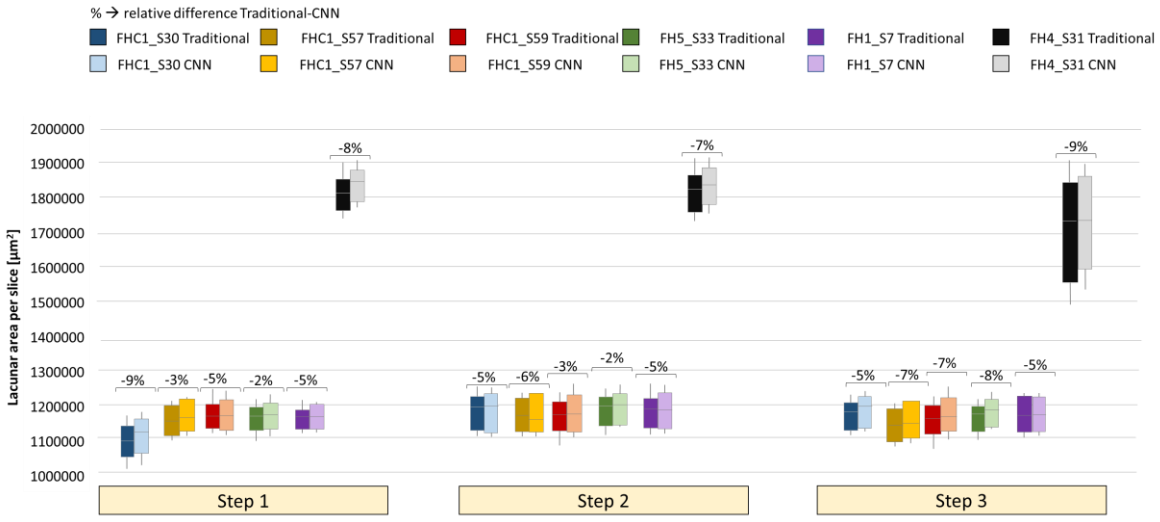


Figure 10: Values of lacunar area per slice measured both by traditional and CNN-based detection. The error bars refer to the standard deviation of the analysed images for each sample at each compression level, considering the tiling process too. The percentage values are related to the percentage relative difference between traditional and CNN-based crack recognition (Eq.2).

Additionally, an exemplificative visual output of the network is schematized in Figure 11: the CNN detection succeeds in highlighting lacunae for both high quality (Figure 11a) and artefacts-affected images (Figure 11b). The presence of a lower number of lacunae in Figure 11b, indeed, is linked just to the presence of possible movement-related artefacts during synchrotron acquisition and not to a specific biomechanical reason.

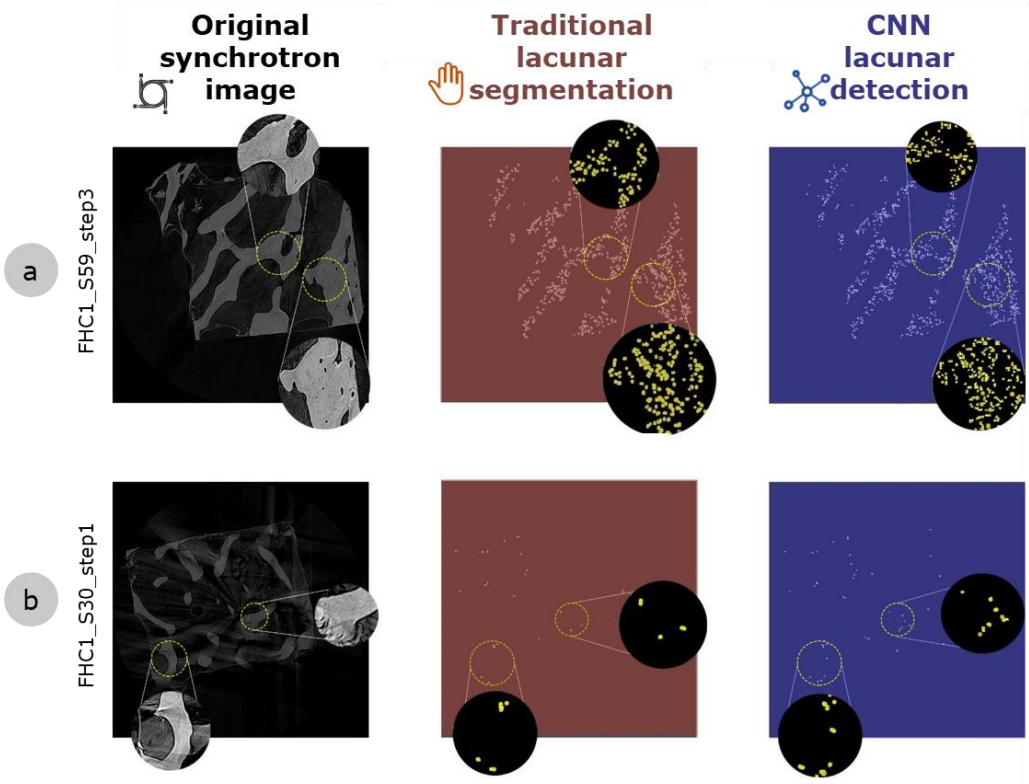


Figure 11: Lacunae detection for FHC1_S30_Z0_step0 and step3 (left) through the network (center) and via algorithm segmentation detailed in paragraph 2.4 (right).

The ability of the CNN in detecting micro-cracks and lacunae could be well visually perceived when moving from the 2D images to the 3D reconstruction of such features (Figure 12). This will allow to highlight the interactions between micro-porosities and micro-fractures, specifically when visualizing the 3D reconstructions at different compression steps.

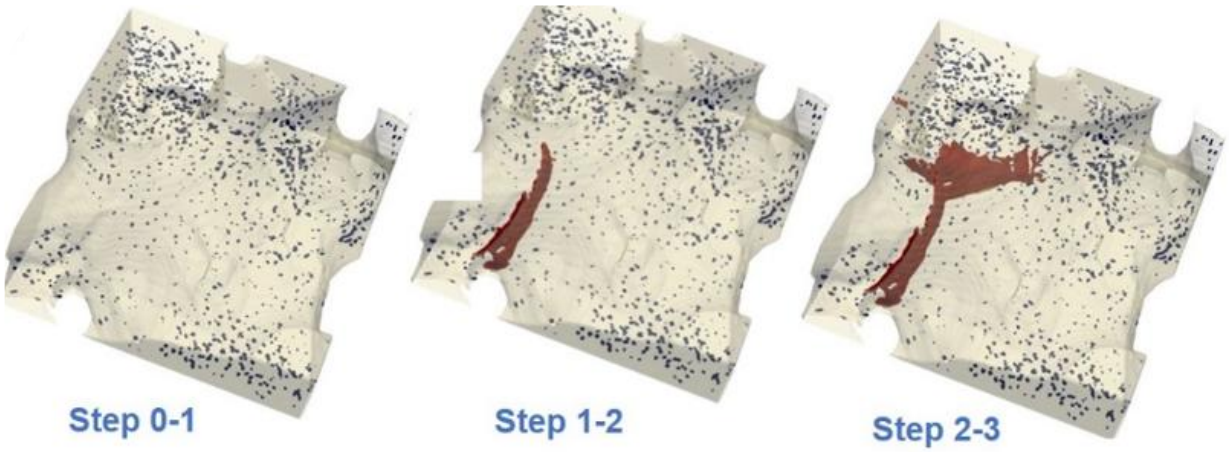


Figure 12: 3D reconstruction of an exemplificative sub-region of FHC1_S59 sample at increasing applied compression (from the left to the right). Lacunae are highlighted in blue, while cracks are depicted in red.

## 4. Discussion

This work aims at shedding some light on the complexity of human bone microdamage by combining the extremely high-resolution and phase contrast provided by the synchrotron light with the automatic detection of bone micro-features given by neural network implementation. More in depth, the adopted methodological approach allows to test femoral samples directly inside the synchrotron and to capture damage progression at different time intervals. However, multiple images acquired during the testing phase induce a considerable computational cost for the reconstruction and post-processing of the two-dimensional images, making the identification of micro-cracks extremely complex and time demanding. For this reason, the synergy with the implemented CNNs permits to separately identify micro-cracks and lacunae at progressive compression steps. The context we are exploring in the present manuscript is a novelty with respect to the state-of-the-art application of CNN to biomedical images. Indeed, common applicative cases deepen the use of CNN for i.e. tumor detection and the goal in those scenarios is to assist less experienced surgeons or clinical doctors in the diagnostic process. In the framework we are analyzing, the difficulty is not the mere detection of micro-scale bone features but the enormous time-consuming process that the manual identification requires. This is specifically linked to the characteristics of synchrotron images, that are the key to perceive different micro-crack patterns, appearing in different regions and sub-regions and in both healthy and osteoporotic bones. Tiling pre-processing stage is extremely beneficial since both cracks and lacunae have small dimensions with respect to the whole specimen section, and it would be unfeasible for a Neural Network to detect many of them without recurring to a zooming and tiling stage (especially for lacunae). Hence, self-similarity of bones at the microscale has been crucial for obtaining an optimal balance between time-demanding high-resolution data acquisition process and data variability in terms of zones, sub-zones and physio-pathological condition.

More in detail, it is interesting to point out that the CNN succeed in identifying cracks at both step 1 and 2 with an accuracy that is higher with respect to manual detection. This is particularly relevant, since, generally, especially at the first compression level, just tiny micro-cracks appear (Figure 8b). These cracks are often starting from external discontinuities that are localized at trabecular margins and visibly progress just when the applied displacement is increased, typically at step 2. The success rate of the CNN in detecting micro-cracks at step 1 is a crucial point for the comprehension of damage initiation sites, that are often not captured by manual analysis of high-resolution images. Conversely, the relative difference between Manual and CNN detection assumes a positive sign in the majority of step 3 samples, reducing the accuracy, as reported in Figure 9. At this stage, cracks are more difficultly highlighted by the implemented AI-based algorithm. This is due to the fact that an advancement in the compression steps generate a collapse in bone structure: cracks cross the trabecula (Figure 8d) and bone fragments separate from the sample body. Indeed, micro-cracks coalesce to generate meso-scale fracture, that involves the whole trabecula. For this reason, the positive sign of the relative difference between manual and CNN segmentation value does not result in a lack of the CNN in detecting bone fractures, since no micro-cracks are present in full bone.

Additionally, observing the trend of cracks area recognition for samples coming from the same femoral head (i.e. FHC1_S30, FHC1_S57 and FHC1_S59), a statistical significative variability could be detected. Indeed, micro-crack area varies from 30000 $\mu m^2$ to around 75000 $\mu m^2$ at the first compression steps and from 130000 $\mu m^2$ to around 210000 $\mu m^2$ at the third compression step. This results confirms previous observations [52] that identify a significative intra-patient variability. This depends from zones of samples' origin: indeed, the three chosen samples are selected from three different sub-zones (proximal, medial and distal with respect to the pelvis). These sub-regions result in different mechanical characteristics, that consequently reflect in different micro-cracks numbers and micro-damage progression patterns. Additionally, it could be noted that criticalities in the synchrotron scanning process, resulting in an unrebuildable output image, affect both manual and CNN-based detection (Figure 8c and 11b).

As regards the comparison between the cracks detected in healthy samples and in the osteoporotic one, it is notable to point out that, even at the third compression step where there is an evident lack in highlighting meso-scale cracks by means of the CNN, the average mismatch between manual and CNN-based recognition for healthy subjects is about +0.4%. This value increases to 12% in presence of an osteoporotic sample. The reason beyond this trend should be checked in the reduced mechanical characteristics of the osteoporotic FH4. Indeed, osteoporotic bone at step 3 is completely fractured in sub-parts, evidently increasing the presence of larger scale cracks. This implies that with the used training images, where mainly healthy samples are chosen, it is not possible to achieve good performances in terms of prediction of peculiar behaviours in osteoporotic samples. On the contrary, from a diagnostical point of view, the higher accuracy in detecting initial stages of cracks finds an increased interest, since meso-and especially macro-scale cracks already represent a critical stage for the subject and, in the specific macro-scale case, fractures could be highlighted by means of traditional clinical images.

Concerning lacunae detection, CNN identification obtains optimal results, reducing the computational cost of manual segmentation of those micrometric elliptical features. As mentioned for cracks identification, the CNN shows improved results with respect to traditional approaches even in presence of artifacts-affected images (Figure 11b).

Although it is possible to build a multiple-feature detection network, returning a multi-label mask considering both lacunae and cracks detection, we preferred to keep the two outputs separated in order to independently check the convergence of the network for lacunae and cracks detection. However, the choice for PNG images allow for implementing a unique CNN for bone multi-feature detection. This is a crucial aspect for assessing the interaction between micro-cracks and lacunae, whose features are currently separately considered.

To the best of authors' knowledge, this is the first work that combines the power of neural networks with the use of synchrotron images of human bones, so a comparison with past researches in terms of NN accuracy would be ineffective. An interesting comparison could be however performed with the recent research of Chin-yun Shen et al. [34], whose neural networks are based on synchrotron scans of cortical and trabecular animal bones. In the present work, we improved the performance of the Convolutional Neural Network approach in terms of accuracy of nearly 5%. More in depth, if we compare the performance of the CNN on the trabecular bone, that is the focus of our work, since we tested human trabecular bones from femoral heads, the improvement in the accuracy is of 9%.

The high accuracy obtained in CNN micro-crack and lacunae detection suggests the possibility to translate two-dimensional micro-scale findings to three-dimensional considerations, in order to have insights on the spatial distribution of micro-damage in comparison with the lacunar network arrangement (Figure 12). This would be extremely beneficial to evaluate three-dimensional interactions among cracks and lacunae, that are currently limited to bidimensional observations. This perspective will pave the way for computational models able to predict damage initiation and evolution up to the final failure, identifying a micro-scale fragility index. Micro-scale findings could be then related to macro-scale images, commonly implied in the clinical practice, adding valuable insights to perform early diagnosis of deteriorated bone architecture and pathology-induced fractures.

## 5. Conclusions

The main outcomes and novelties of the work are summarized as follows:

- for the first time, a CNN is used for the detection of bone micro-architecture and bone micro-damage progression for synchrotron images of human samples;

- healthy and osteoporotic samples are tested under micro-compression and progressive images are obtained at the micrometric resolution;
- specifically, a VGG16 Neural Network has been implemented and trained for the identification of lacunae and micro-cracks;
- convergence assessments in terms of accuracy and meanIoU for both networks are presented. In particular, the implemented CNN achieves performant values of 0.8743 meanIoU for micro-cracks detection and 0.8751 meanIoU for lacunae recognition;
- CNNs achieve more than 0.99 level of accuracy in cracks detection; this result is particularly encouraging, also considering the complexity of the scanning procedure and the highlighted criticalities when micro-cracks coalesce into larger-scale meso-cracks;
- the work offers a unique potentiality to detect automatically cracks and lacunae: this will pave the way for further quantitative observations on bone micro-scale architecture and damage. In particular, the isolation of cracks will allow to reconstruct them three-dimensionally and to build the basis for computational models focused on bone micro-level.

## Bibliography


1. Burge, R.; Dawson-Hughes, B.; Solomon, D.H.; Wong, J.B.; King, A.; Tosteson, A. Incidence and economic burden of osteoporosis-related fractures in the United States, 2005-2025. *J. Bone Miner. Res.* **2007**, *22*, 465–475, doi:10.1359/jbmr.061113.

2. Nalla, R.K.; Kruzic, J.J.; Kinney, J.H.; Balooch, M.; Ager, J.W.; Ritchie, R.O. Role of microstructure in the aging-related deterioration of the toughness of human cortical bone. *Mater. Sci. Eng. C* **2006**, *26*, 1251–1260, doi:10.1016/j.msec.2005.08.021.

3. Goltzman, D. The Aging Skeleton. In *Advances in Experimental Medicine and Biology*; Springer New York LLC, 2019; Vol. 1164, pp. 153–160.

4. Schaffler, M.B.; Burr, D.B. Stiffness of compact bone: Effects of porosity and density. *J. Biomech.* **1988**, *21*, 13–16, doi:10.1016/0021-9290(88)90186-8.

5. Sabet, F.A.; Najafi, A.R.; Hamed, E.; Jasiuk, I. Modelling of bone fracture and strength at different length scales: A review. *Interface Focus* 2016, *6*.

6. Jerome, C.; Hoch, B. Skeletal System. *Comp. Anat. Histol.* 2012, 53–70.

7. Meyers, M.A.; Chen, P.Y. Biological materials science: Biological materials, bioinspired materials, and biomaterials. *Biol. Mater. Sci. Biol. Mater. Bioinspired Mater. Biomater.* 2014, 1–628.

8. Ratner, B.D.; Hoffman, A.S.; Schoen, F.J.; Lemons, J.E. *Biomaterials Science: An Introduction to Materials: Third Edition*; Elsevier Inc., 2013; ISBN 9780123746269.

9. Qiu, S.; Rao, D.S.; Fyhrie, D.P.; Palnitkar, S.; Parfitt, A.M. The morphological association between microcracks and osteocyte lacunae in human cortical bone. *Bone* **2005**, *37*, 10–15, doi:10.1016/j.bone.2005.01.023.

10. Portier, H.; Jaffré, C.; Kewish, C.M.; Chappard, C.; Pallu, S. New insights in osteocyte imaging by synchrotron radiation. *J. Spectr. Imaging* **2020**, *9*, doi:10.1255/jsi.2020.a3.

11. Carter, Y.; Thomas, C.D.L.; Clement, J.G.; Cooper, D.M.L. Femoral osteocyte lacunar density, volume and morphology in women across the lifespan. *J. Struct. Biol.* **2013**, *183*, 519–526, doi:10.1016/J.JSB.2013.07.004.

12. McCreadie, B.R.; Hollister, S.J.; Schaffler, M.B.; Goldstein, S.A. Osteocyte lacuna size and shape in women with and without osteoporotic fracture. *J. Biomech.* **2004**, *37*, 563–572, doi:10.1016/S0021-9290(03)00287-2.

13. Zimmermann, E.A.; Busse, B.; Ritchie, R.O. The fracture mechanics of human bone: influence of disease and treatment. *Bonekey Rep.* **2015**, *4*, doi:10.1038/bonekey.2015.112.

14. CURREY, J.D. Stress Concentrations in Bone. *J. Cell Sci.* **1962**, *s3-103*, 111–133, doi:10.1242/jcs.s3-103.61.111.

15. Burr, D.B. Stress concentrations and bone microdamage: John Currey's contributions to understanding the initiation and arrest of cracks in bone. *Bone* **2019**, *127*, 517–525, doi:10.1016/j.bone.2019.07.015.

16. Reilly, G.C. Observations of microdamage around osteocyte lacunae in bone. *J. Biomech.* **2000**, *33*, 1131–1134, doi:10.1016/S0021-9290(00)00090-7.

17. Prendergast, P.J.; Huiskes, R. Microdamage and osteocyte-lacuna strain in bone: A microstructural finite element analysis. *J. Biomech. Eng.* **1996**, *118*, 240–246, doi:10.1115/1.2795966.

18. Libonati, F.; Nair, A.K.; Vergani, L.; Buehler, M.J. Fracture mechanics of hydroxyapatite single crystals under geometric confinement. *J. Mech. Behav. Biomed. Mater.* **2013**, *20*, 184–191, doi:10.1016/j.jmbbm.2012.12.005.

19. Schneider, P.; Voide, R.; Stampanoni, M.; Donahue, L.R.; Müller, R. The importance of the intracortical canal network for murine bone mechanics. *Bone* **2013**, *53*, 120–128, doi:10.1016/j.bone.2012.11.024.

20. Awaji, H.; Ebisudani, M.; Choi, S.M.; Ohashi, T. Crack deflection toughening mechanism in brittle materials. In *ASTM Special Technical Publication*; ASTM International, 2002; pp. 137–151.

21. Reznikov, N.; Shahar, R.; Weiner, S. Bone hierarchical structure in three dimensions. *Acta Biomater.* **2014**, *10*, 3815–3826, doi:10.1016/j.actbio.2014.05.024.

22. Zioupos, P. Recent developments in the study of failure of solid biomaterials and bone: "Fracture" and "pre-fracture" toughness. *Mater. Sci. Eng. C* **1998**, *6*, 33–40, doi:10.1016/S0928-4931(98)00033-2.

23. Buccino, F.; Colombo, C.; Duarte, D.H.L.; Rinaudo, L.; Ulivieri, F.M.; Vergani, L.M. 2D and 3D numerical models to evaluate trabecular bone damage. *Med. Biol. Eng. Comput.* **2021**, *59*, 2139–2152, doi:10.1007/s11517-021-02422-x.

24. Ma, Y.; Luo, Y. Bone fracture detection through the two-stage system of Crack-Sensitive Convolutional Neural Network. In *Informatics in Medicine Unlocked*; 2021; Vol. 22.

25. Shaikh, A.B.; Sarim, M.; Raffat, K.; Ahsan, K.; Nadeem, A.; Siddiq, M. Artificial Neural Network: A Tool for Diagnosing Osteoporosis. *Res. J. Recent Sci.* **2014**, *3*, 87–91.

26. Pradhan, N.; Singh Dhaka, V.; Chaudhary, H. Classification of Human Bones Using Deep Convolutional Neural Network. In Proceedings of the IOP Conference Series: Materials Science and Engineering; 2019; Vol. 594.

27. Liu, Q.; Cui, X.; Chou, Y.C.; Abbod, M.F.; Lin, J.; Shieh, J.S. Ensemble artificial neural networks applied to predict the key risk factors of hip bone fracture for elders. In *Biomedical Signal Processing and Control*; 2015; Vol. 21, pp. 146–156.

28. Yang, A.Y.; Cheng, L.; Shimaponda-Nawa, M.; Zhu, H.Y. Long-Bone Fracture Detection using Artificial Neural Networks based on Line Features of X-ray Images. In Proceedings of the 2019 IEEE Symposium Series on Computational Intelligence, SSCI 2019; 2019; pp. 2595–2602.

29. Rehman, F.; Ali Shah, S.I.; Riaz, M.N.; Gilani, S.O.; R, F. A Region-Based Deep Level Set Formulation for Vertebral Bone Segmentation of Osteoporotic Fractures. *J. Digit. Imaging*

**2020**, *33*, 191–203, doi:10.1007/s10278-019-00216-0.

30. Wang, S.; Shen, Y.; Zeng, D.; Hu, Y. Bone age assessment using convolutional neural networks. In Proceedings of the 2018 International Conference on Artificial Intelligence and Big Data, ICAIBD 2018; 2018; pp. 175–178.

31. Villa, M.; Dardenne, G.; Nasan, M.; Letissier, H.; Hamitouche, C.; Stindel, E. FCN-based approach for the automatic segmentation of bone surfaces in ultrasound images. In Proceedings of the International Journal of Computer Assisted Radiology and Surgery; Beijing, China, 2018; Vol. 13, pp. 1707–1716.

32. Hemke, R.; Buckless, C.G.; Tsao, A.; Wang, B.; Torriani, M. Deep learning for automated segmentation of pelvic muscles, fat, and bone from CT studies for body composition assessment. *Skeletal Radiol.* **2020**, *49*, 387–395, doi:10.1007/s00256-019-03289-8.

33. Deniz, C.M.; Xiang, S.; Hallyburton, R.S.; Welbeck, A.; Babb, J.S.; Honig, S.; Cho, K.; Chang, G. Segmentation of the Proximal Femur from MR Images using Deep Convolutional Neural Networks. *Sci. Rep.* **2018**, *8*, 16485, doi:10.1038/s41598-018-34817-6.

34. Shen, S.C. yun; Peña Fernández, M.; Tozzi, G.; Buehler, M.J. Deep learning approach to assess damage mechanics of bone tissue. *J. Mech. Behav. Biomed. Mater.* 2021, *123*, 104761.

35. Lindsey, R.; Daluiski, A.; Chopra, S.; Lachapelle, A.; Mozer, M.; Sicular, S.; Hanel, D.; Gardner, M.; Gupta, A.; Hotchkiss, R.; et al. Deep neural network improves fracture detection by clinicians. In Proceedings of the Proceedings of the National Academy of Sciences of the United States of America; 2018; Vol. 115, pp. 11591–11596.

36. Messina, C.; Piodi, L.P.; Grossi, E.; Eller-Vainicher, C.; Bianchi, M.L.; Ortolani, S.; Stefano, M. Di; Rinaudo, L.; Sconfienza, L.M.; Ulivieri, F.M. Artificial neural network analysis of bone quality DXA parameters response to teriparatide in fractured osteoporotic patients. In *PLoS ONE*; 2020; Vol. 15.

37. Frank, M.; Fischer, J.T.; Thurner, P.J. Microdamage formation in individual bovine trabeculae during fatigue testing. *J. Biomech.* **2021**, *115*, 110131, doi:10.1016/j.jbiomech.2020.110131.

38. Larrue, A.; Rattner, A.; Peter, Z.A.; Olivier, C.; Laroche, N.; Vico, L.; Peyrin, F. Synchrotron radiation micro-CT at the Micrometer scale for the analysis of the three-dimensional morphology of microcracks in human trabecular bone. *PLoS One* **2011**, *6*, e21297, doi:10.1371/journal.pone.0021297.

39. Wentzell, S.; Nesbitt, R.S.; Macione, J.; Kotha, S. Measurement of lacunar bone strains and crack formation during tensile loading by digital volume correlation of second harmonic generation images. *J. Mech. Behav. Biomed. Mater.* **2016**, *60*, 148–156, doi:10.1016/j.jmbbm.2015.12.027.

40. Voide, R.; Schneider, P.; Stauber, M.; Wyss, P.; Stampanoni, M.; Sennhauser, U.; van Lenthe, G.H.; Müller, R. Time-lapsed assessment of microcrack initiation and propagation in murine cortical bone at submicrometer resolution. *Bone* **2009**, *45*, 164–173, doi:10.1016/j.bone.2009.04.248.

41. Voide, R.; Schneider, P.; Stauber, M.; van Lenthe, G.H.H.; Stampanoni, M.; Müller, R.; Stauber, M.; Schneider, P.; van Lenthe, G.H.H.; Voide, R.; et al. The importance of murine cortical bone microstructure for microcrack initiation and propagation. *Bone* **2011**, *49*, 1186–1193, doi:10.1016/j.bone.2011.08.011.

42. Buccino, F.; Colombo, C.; Vergani, L.M. A review on multiscale bone damage: From the clinical to the research perspective. *Materials (Basel).* **2021**, *14*, 1–25, doi:10.3390/ma14051240.

43. Buccino, F.; Bagherifard, S.; D'Amico, L.; Zagra, L.; Banfi, G.; Tromba, G.; Vergani, L.M.

Assessing the intimate mechanobiological link between human bone micro-scale trabecular architecture and micro-damages. *Eng. Fract. Mech.* **2022**, *270*, 108582, doi:10.1016/J.ENGFRACMECH.2022.108582.

44. Larson, S.M.; Fishbein, D.; Finn, R.D.; Baldwin, G.P. Assessment of goals and obstacles in data acquisition and analysis from emission tomography. Report of a series of international workshops. II. Quality control issues. *J. Cereb. Blood Flow Metab.* **1987**, *7*, S1–S3, doi:10.1038/JCBFM.1987.25.

45. Dudak, J.; Zemlicka, J.; Karch, J.; Patzelt, M.; Mrzilkova, J.; Zach, P.; Hermanova, Z.; Kvacek, J.; Krejci, F. High-contrast X-ray micro-radiography and micro-CT of ex-vivo soft tissue murine organs utilizing ethanol fixation and large area photon-counting detector. *Sci. Rep.* **2016**, *6*, 1–9, doi:10.1038/srep30385.

46. Mehta, N.; Braun, P.X.; Gendelman, I.; Alibhai, A.Y.; Arya, M.; Duker, J.S.; Waheed, N.K. Repeatability of binarization thresholding methods for optical coherence tomography angiography image quantification. *Sci. Rep.* **2020**, *10*, 1–11, doi:10.1038/s41598-020-72358-z.

47. Chu, Z.; Gregori, G.; Rosenfeld, P.J.; Wang, R.K. Quantification of Choriocapillaris with Optical Coherence Tomography Angiography: A Comparison Study. *Am. J. Ophthalmol.* **2019**, *208*, 111–123, doi:10.1016/j.ajo.2019.07.003.

48. Onur, T.O. An application of filtered back projection method for computed tomography images. *Int. Rev. Appl. Sci. Eng.* **2021**, *12*, 194–200, doi:10.1556/1848.2021.00231.

49. Zhang, H.; Li, L.; Qiao, K.; Wang, L.; Yan, B.; Li, L.; Hu, G. Image Prediction for Limited-angle Tomography via Deep Learning with Convolutional Neural Network. **2016**.

50. Waarsing, J.H.; Day, J.S.; Van Der Linden, J.C.; Ederveen, A.G.; Spanjers, C.; De Clerck, N.; Sasov, A.; Verhaar, J.A.N.; Weinans, H. Detecting and tracking local changes in the tibiae of individual rats: A novel method to analyse longitudinal in vivo micro-CT data. *Bone* **2004**, *34*, 163–169, doi:10.1016/j.bone.2003.08.012.

51. Kiziltaş, G.; Papila, M.; Yilmaz, B.; Bilge, K. Challenges in Micro-CT Characterization of Composites. *Micro-computed Tomogr. Med. Eng.* **2020**, 225–246, doi:10.1007/978-3-030-16641-0_14.

52. Buccino, F.; Zagra, L.; Savadori, P.; Colombo, C.; Grossi, G.; Banfi, G.; Vergani, L. Mapping Local Mechanical Properties of Human Healthy and Osteoporotic Femoral Heads. *SSRN Electron. J.* **2021**, doi:10.2139/ssrn.3878345.

53. Barth, H.D.; Launey, M.E.; MacDowell, A.A.; Ager, J.W.; Ritchie, R.O. On the effect of X-ray irradiation on the deformation and fracture behavior of human cortical bone. *Bone* **2010**, *46*, 1475–1485, doi:10.1016/j.bone.2010.02.025.

54. Paganin, D.; Mayo, S.C.; Gureyev, T.E.; Miller, P.R.; Wilkins, S.W. Simultaneous phase and amplitude extraction from a single defocused image of a homogeneous object. *J. Microsc.* **2002**, *206*, 33–40, doi:10.1046/j.1365-2818.2002.01010.x.

55. GitHub - ElettraSciComp/STP-Gui: SYRMEP Tomo Project (.NET Graphical User Interface for STP-Core) Available online: https://github.com/ElettraSciComp/STP-Gui (accessed on Sep 29, 2021).

56. Brun, F.; Massimi, L.; Fratini, M.; Dreossi, D.; Billé, F.; Accardo, A.; Pugliese, R.; Cedola, A. SYRMEP Tomo Project: a graphical user interface for customizing CT reconstruction workflows. *Adv. Struct. Chem. Imaging* **2017**, *3*, doi:10.1186/S40679-016-0036-8.

57. Goff, E.; Buccino, F.; Bregoli, C.; McKinley, J.P.; Aeppli, B.; Recker, R.R.; Shane, E.; Cohen, A.; Kuhn, G.; Müller, R. Large-scale quantification of human osteocyte lacunar morphological biomarkers as assessed by ultra-high-resolution desktop micro-computed tomography. *Bone*

**2021**, *152*, doi:10.1016/j.bone.2021.116094.

58. McCreadie, B.R.; Hollister, S.J.; Schaffler, M.B.; Goldstein, S.A. Osteocyte lacuna size and shape in women with and without osteoporotic fracture. *J. Biomech.* **2004**, *37*, 563–572, doi:10.1016/S0021-9290(03)00287-2.

59. Bach-Gansmo, F.L.; Brüel, A.; Jensen, M.V.; Ebbesen, E.N.; Birkedal, H.; Thomsen, J.S. Osteocyte lacunar properties and cortical microstructure in human iliac crest as a function of age and sex. *Bone* **2016**, *91*, 11–19, doi:10.1016/J.BONE.2016.07.003.

60. Akhter, M.P.; Kimmel, D.B.; Lappe, J.M.; Recker, R.R. Effect of Macroanatomic Bone Type and Estrogen Loss on Osteocyte Lacunar Properties in Healthy Adult Women. *Calcif. Tissue Int.* **2017**, *100*, 619–630, doi:10.1007/S00223-017-0247-6.

61. Larrue, A.; Rattner, A.; Peter, Z.A.; Olivier, C.; Laroche, N.; Vico, L.; Peyrin, F. Synchrotron Radiation Micro-CT at the Micrometer Scale for the Analysis of the Three-Dimensional Morphology of Microcracks in Human Trabecular Bone. *PLoS One* **2011**, *6*, e21297, doi:10.1371/JOURNAL.PONE.0021297.

62. Simonyan, K.; Zisserman, A. Very deep convolutional networks for large-scale image recognition. *3rd Int. Conf. Learn. Represent. ICLR 2015 - Conf. Track Proc.* **2015**.

63. Kingma, D.P.; Ba, J.L. Adam: A method for stochastic optimization. *3rd Int. Conf. Learn. Represent. ICLR 2015 - Conf. Track Proc.* **2015**.

64. Chen, P.; Gao, L.; Shi, X.; Allen, K.; Yang, L. Fully automatic knee osteoarthritis severity grading using deep neural networks with a novel ordinal loss. *Comput. Med. Imaging Graph.* **2019**, *75*, 84–92, doi:10.1016/J.COMPMEDIMAG.2019.06.002.

65. Atanbori, J.; French, A.P.; Pridmore, T.P. Towards infield, live plant phenotyping using a reduced-parameter CNN. *Mach. Vis. Appl.* **2020**, *31*, 1–14, doi:10.1007/s00138-019-01051-7.