Data-driven image mechanics (D^{2}IM): A deep learning approach to predict displacement and strain fields from undeformed X-ray tomography images – Evaluation of bone mechanics

Peter Soar [a], Marco Palanca [b], Enrico Dall'Ara [c,d], Gianluca Tozzi [e,*]

[a] Centre for Advanced Simulation and Modelling, School of Computing and Mathematical Sciences, University of Greenwich, United Kingdom
[b] Department of Industrial Engineering, Alma Mater Studiorum – University of Bologna, Italy
[c] Division of Clinical Medicine, University of Sheffield, UK
[d] Insigneo Institute, University of Sheffield, UK
[e] Centre for Advanced Manufacturing and Materials, School of Engineering, University of Greenwich, United Kingdom

ARTICLE INFO

ABSTRACT

The recent advent of deep learning (DL) has enabled data-driven models to pave the way for the full exploitation of rich image datasets from which physics can be learnt. Here we propose a novel data-driven image mechanics (D^{2}IM) approach that learns from digital volume correlation (DVC) displacement fields of vertebrae, predicting displacement and strain fields from undeformed X-ray computed tomography (XCT) images. D^{2}IM successfully predicted the displacements in all directions, particularly in the cranio-caudal direction of the vertebra, where high correlation (R^2=0.94) and generally minimal errors were obtained compared to the measured displacements. The predicted axial strain field in the cranio-caudal direction of the vertebra was also consistent in distribution with the measured one, displaying generally reduced errors in the regions within the vertebral body. The application of D^{2}IM to lower resolution imaging in initial testing provides promising results indicating the future viability of integrating this technology into a clinical setting. This is the first study using experimental full-field measurements on bone structures from DVC to inform DL-based models such as D^{2}IM, which represents a major contribution the prediction of displacement and strain fields based only on the greyscale content of undeformed XCT images. In future, D^{2}IM will incorporate a range of biological structures and loading scenarios for accurate prediction of physical fields, aiming at clinical translation for improved diagnostics.
Data Availability: Code for preparing dataset, training D^{2}IM model and visualising/analysing results has been hosted on GitHub: https://github.com/PeterSoar/D2IM_Prototype
The dataset used for this study can be found on Figshare: https://doi.org/10.6084/m9.figshare.25404220.v1

1. Introduction

The intricate biomechanical properties of musculoskeletal tissues and their response to mechanical loading are crucial for understanding disease effects and optimising treatments. A prime example is bone fracture risk linked to the tissue's ability to withstand crack propagation, which is closely associated with its distinctive deformation behaviour. Experimental techniques such as digital volume correlation (DVC) are currently considered state-of-the-art in the mechanical characterization of bone, from organ to tissue level [1]. DVC allows for full-field displacement and strain measurements in the material volume by correlating grayscale features from three-dimensional images, typically obtained via high-resolution X-ray computed tomography (XCT), before and after the application of mechanical loading in a process known as in situ XCT mechanics [2]. Traditional in situ XCT mechanics is performed in a stepwise fashion (i.e. time-lapsed testing), with each loading phase followed by a holding period of 15–30 min for full tomographic acquisition to reduce the impact of moving artifacts due to stress relaxation, which is crucial for biological tissues due to their viscoelastic behaviour [3]. Thus, in situ experiments can be considerably time consuming as the XCT imaging (i.e. higher resolution) and mechanical (i.e. increased number of in situ steps) requirements are more demanding, with experiment duration for one specimen that can vary from a few hours to more than 24 h [4,5]. This can result in extended exposure to X-rays,

which is proven to have a substantial impact on the structural and mechanical properties of musculoskeletal tissues both in high-flux synchrotron [6] and laboratory XCT setups [4]. In addition, experimental tests to characterise bone using in situ XCT mechanics and DVC are typically conducted with a limited sample size of 2–3 specimens per bone type due to tissue availability, particularly in cases of in vivo treatment [7]. This makes the analysis more qualitative and valid on a case-to-case basis but difficult to generalise within the same cohort, even in the case of tissue samples of similar size, extracted from the same anatomical site and tested using the same protocol. Whilst the above aspects are characteristic of any XCT-based DVC analysis and may be mitigated by appropriate strategies, there is a need to optimise quality and number of experiments for a given bone type/structure making an efficient use of acquired images and associated full-field measurements to predict physical fields under selected loading conditions. But as DVC algorithms rely solely on the visualised greyscale features in the input images, for example bone trabeculae resolved in the acquired XCT images, they lack direct awareness of the object's material properties. In this context, experimental fields from DVC have been extensively used to validate finite element (FE) models derived from XCT images, which can then predict mechanical properties in bone, such as vertebral bodies, both with [8] and without lesions [9]. Despite the popularity of validation and use of FE models based on DVC, this procedure is technically challenging and based on some material property assumptions varying in accordance with the specific loading regime to simulate (i.e. linear elastic).

The emergence of machine learning (ML), particularly deep learning (DL), has ushered in a new era for the swift resolution of intricate tasks such as these. Different from traditional ML with feature extractors, DL essentially belongs to a class of data-driven end-to-end models, which has achieved great success in different bioengineering areas such as medical image segmentation and classification [10]. Convolutional neural networks (CNNs) are probably the most popular class of DL models employed in imaging as they possess the ability to learn complex features by extracting visual information automatically using combinations of series of transformations in the model architecture. The typical architecture of CNN has a multi-layer feed-forward network with an input layer, hidden layers including convolutional layers projecting a series of image filters onto the input image and fully connected dense layers, with an output layer where the predictions will be extracted. CNNs are generally robust with low complexity and easy to train, where the network learns throughout the optimization process with a reduced number of parameters [11]. CNNs, have already demonstrated prowess in classifying stages of bone tissue deformation leading to fractures as well as segmenting cracks, employing both high-resolution synchrotron [12,13] and laboratory XCT [14] in situ mechanics. Recently, DL has been integrated in two-dimensional digital image correlation (DIC) [15–18] and DVC [10]. In both cases, DL has shown remarkable promise by significantly reducing computational complexity, thereby enhancing efficiency in analyses. Additionally, a synergistic approach has been proposed, coupling DL with the traditional cross-correlation method from particle image velocimetry (PIV), to optimize and refine a coarse velocity field, yielding super-resolution calculations [19]. This area holds immense potential for enhancing measurements and advancing models for comprehending and predicting the mechanics of musculoskeletal tissues. Recent advances have seen physics-informed neural networks leveraged to solve problems generally formulated as partial differential equations by predicting full-field data for many processes, notably including crack propagation [20,21] and mechanical fields such as displacements, stress and strain within a structure [21–25]. In this regard, DL models equipped with the capacity to predict physical fields, such as stress or strain, directly from simple images encapsulating geometry and microstructure information that fully encodes material composition and boundary conditions have emerged [26,27]. AI-based frameworks have been developed to predict comprehensive strain and stress fields using partial data, enabling the inverse translation from

mechanical fields to composite microstructures [28]. However, it is important to note that both DL-based approaches, whether aimed at improving the efficiency of DIC/DVC analysis or predicting physical fields, have primarily been based on synthetic images or patterns and relatively straightforward geometries and material distributions. Consequently, there has been limited exploration of the potential of these methods to directly operate on the grayscale content of XCT imaging for complex biological structures, such as bone, to predict full field measurements.

This work utilises the published datasets of high-resolution in situ XCT images of intact and artificially lesioned vertebral bone [29] previously used to measure full-field displacements and strains with DVC [30] and validate FE models [8], simulating the mechanical performance of metastatic vertebrae. A larger dataset was derived by augmenting two-dimensional cross-sections of a limited set of three-dimensional XCT tomograms, proposing a new CNN-based approach for data-driven image mechanics (D^{2}IM) to predict displacement and consequently calculate strain fields through a synergistic integration of DVC-measured full-field displacements and deep learning. Preliminary testing of how this framework can be generalised to make predictions using lower-resolutions images was also performed, to demonstrate the potential viability of implementing this technology in a clinical setting where the imaging of patient tissues is acquired at significantly lower visual fidelity than can be achieved using XCT imaging. The findings of this study introduce a major contribution in the prediction of displacement and derived strain fields directly from the grayscale content and texture of undeformed XCT images, setting the scene for advanced image-based prediction of bone deformation and fracture in healthy and pathological conditions.

2. Materials and methods

Porcine vertebrae undergoing in situ stepwise XCT compression were used, and a binary mask was made of the unloaded tomography. These were all used as an input for DVC using the python library SPAM to generate full field displacement data. The 3D images of the unloaded tomogram, mask and three displacement fields were all sliced into 2D images, which could be used as an input and ground truth for D^{2}IM. D^{2}IM is a feed-forward Convolutional Neural Network (CNN) that uses four convolutional stacks with max pooling layers. After being flattened, the model feeds into three dense layers which predict the three displacement components. Strains can be calculated from both the measured and predicted displacement fields for comparison. This entire workflow has been summarised in Fig. 1.

2.1. Tomography Dataset

The data chosen to test this modelling framework are a set of XCT images, with an isotropic voxel size of 39µm, acquired on five porcine thoracic spine segments (two T8-T10, two T10-T12 and one T12-T14) prepared and in situ mechanically tested by Palanca et al. [8,30] with and without artificial lesions to create 10 image pairs overall, which have been made available by the authors on Figshare [29]. This previous work contains the full details of sample preparation, testing and imaging. For each vertebra a tomography has been taken before and after applying a compressive axial load of 6500 N. From the unloaded tomography a binary mask (zero outside the vertebral body and one inside) was created for each image using ImageJ [31], automatically defining the binary threshold based on a slice in the centre of the structure, applying a dilation of 8 voxels, and filling in any remaining internal holes.

2.2. Digital Volume Correlation

Digital volume correlation (DVC) was performed with SPAM (Software for Practical Analysis of Materials), an open-source Python
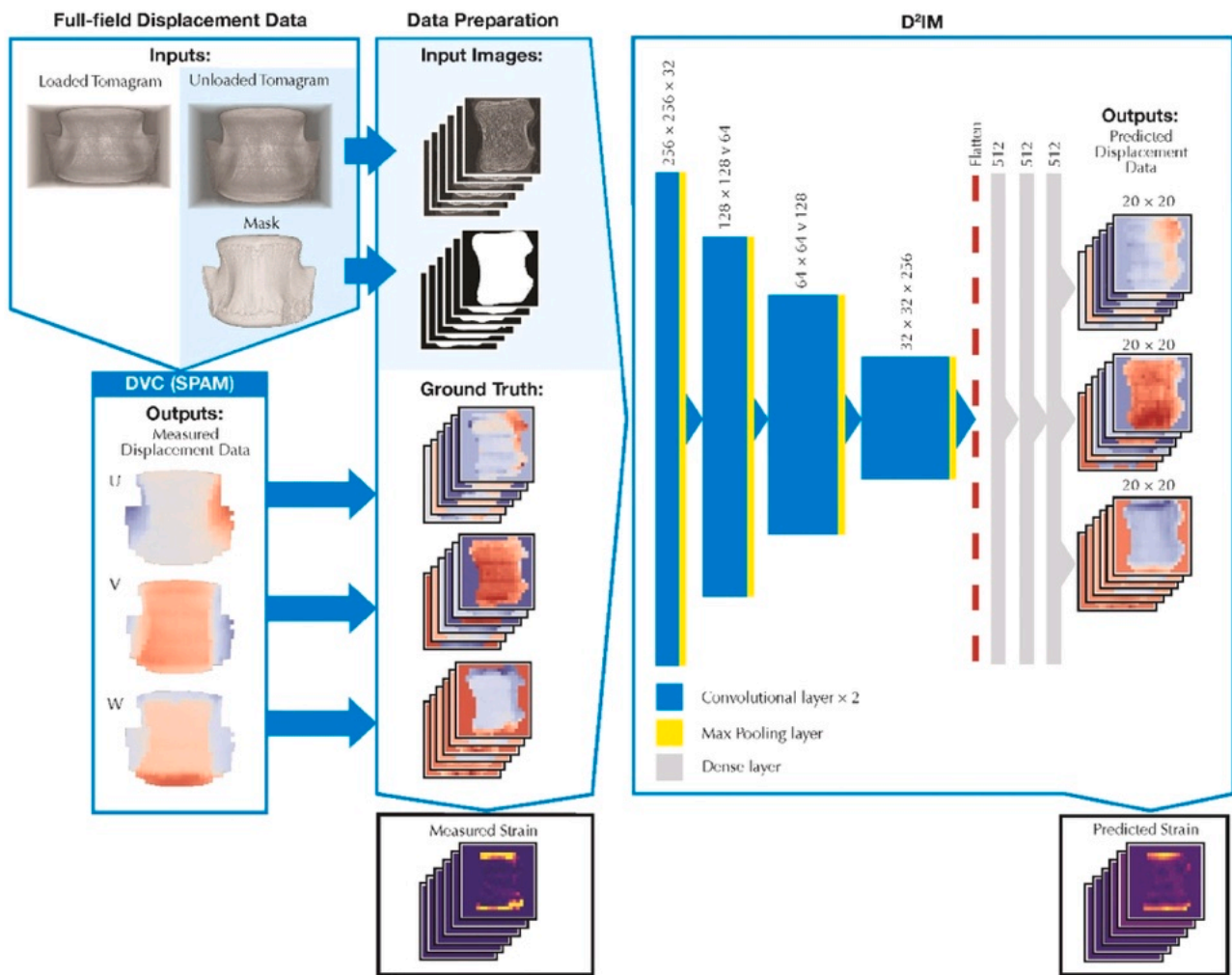
Fig. 1. Workflow for generating 2D sliced dataset and making displacement field predictions using Machine Learning Model D^{2}IM. Strain fields are calculated from the displacements.

correlation software operating on a local approach. SPAM being an open-source Python package was one of the primary reasons for choosing it over other DVC software, as this simplified the integration of DVC into the rest of the data preparation process and allowed more flexibility for integrating with the D^{2}IM in the future. Further details of the operating principles of SPAM have already been reported elsewhere [32,33]. SPAM DVC approach follows a different correlation strategy (i. e. local approach) than the one in Palanca et al. [30] (i.e. BoneDVC, global approach). In brief, local approaches are based on the subdivision of the image into smaller windows (aka sub-volumes) and the spatial correlation of metrics computed in each of the windows of the undeformed and deformed image independently [34]. Global approaches instead are based on the minimisation of the difference of the deformed image and the registered undeformed image when a continuous displacement field is applied [35]. SPAM was used to measure the volumetric displacement fields in vertebrae before and after loading. Using SPAM a non-rigid registration was performed between every loaded and unloaded tomography with a binning of 2 and using the binary mask to identify the region of interest.

The non-rigid registration, raw tomograms and the binary mask are then used in SPAM to measure the full-field displacements in the x (left-right), y (anterior-posterior) and z (cranio-caudal) directions; which will henceforth be referred to as $\bar{u}$, $\bar{v}$ and $\bar{w}$, respectively. When performing DVC, a window size of 50 voxels was selected, consistently with the nodal spacing of 50 voxels used by Palanca et al. [8,30] for the same images. The performance of SPAM (local approach) in measuring

displacements was investigated with a set of consistent benchmark images (voxel size of 39μm) and window/spacing (48 voxels) already used to test BoneDVC for trabecular bone from vertebral bodies [36]. SPAM registered random errors of 3.2μm in x, 2.9μm in y and 2.5μm in z, which are comparable to values (1.73μm in x, 1.57μm in y and 1.36μm in z) obtained by another local DVC software (DaVis-DC) also tested in [36]. The binary mask was again used to identify the region of interest to measure these displacements, where the correlation window had to contain more than 50 % unmasked voxels to be considered an active correlation window by the DVC process. A SPAM filter function was also used on the displacement fields to remove the points with poor DVC convergence, with these points being replaced with a value interpolated from the 27 nearest neighbours.

A comparison was made between the displacement magnitude/distribution measured using SPAM and those obtained by Palanca et al. [30], taking a single intact vertebra case. In Fig. 2 there appears to be a good qualitative match in the displacements registered by both methods on the same vertebra, barring some edge discrepancies from the different masks used. In Table 1 summary statistics for the distribution of the displacement components is reported. According to a Mann-Whitney U test at 95 % significance, the u displacements come from the same distribution however the other components do have significant differences. This is likely due to a combination of the differing DVC methodology and the different masking strategy causing a minor mismatch in the identified structure. Regardless, the same general mechanical behaviour is seemingly still being captured, with less than
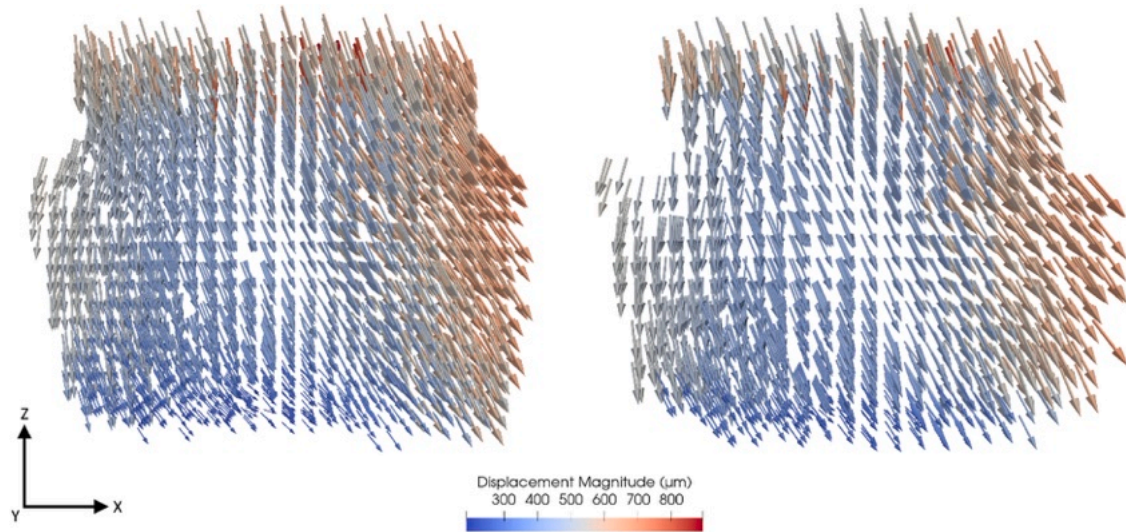
Fig. 2. Displacement field magnitude and distribution for in situ XCT of an intact vertebra from Palanca et al. [30] (left) and SPAM DVC used in this study on the same vertebra (right).

Table 1

Statistics on the distribution of the displacement (μm) components for the DVC results for a single vertebra presented in Fig. 2. Global DVC displacements provided by Palanca et al. [30]. Local DVC displacements measured using SPAM.

DVC Method	Global (BoneDVC)			Local (SPAM)		
Displacement (μm)	u	v	w	u	v	w
Mean	161.2	-94.7	-425.0	161.7	-85.9	-416.1
Median	161.0	-82.5	-431.6	161.0	-75.0	-407.5
Standard Deviation	114.6	51.8	110.4	101.9	39.9	85.7
IQR	91.8	56.3	121.0	50.1	37.3	105.7

10 % variation between the mean measured displacements for both other displacement components. While it is reassuring that SPAM is providing sensible displacement measurements to use as a ground truth for training D^2IM, ultimately the proposed approach can be generalised and applied to any DVC methodology.

2.3. Machine Learning Dataset Preparation

The proposed machine learning (ML) model requires two-dimensional input to output images. The three-dimensional unloaded tomography, binary mask and displacement components for all 10 vertebrae were sliced into sets of 2D images and used as input or target data. The image slicing was performed both in the anterior-posterior and right-left direction for each vertebra, as this would serve to both increase numerosity of the dataset and showcase a wider variety of two-dimensional cases.

To match the number of displacement components for the unloaded scan and the binary mask, only slices corresponding to the central measurement point of each DVC correlation window was kept.

All the images have slightly different dimensions and, as the ML model requires consistent image input and output sizes, they had to be resized. The unloaded image slices from each tomogram were resized to be 256×256 pixels, this is roughly a quarter of the size of the unaltered images which was found to be a good compromise allowing for faster training while still maintaining sufficient resolution.

The binary mask and three displacement components were all resized to be 20×20 image, providing a good approximation across the different scans of average DVC output dimensions. The resizing of the images was done using the *zoom* function in scipy by deforming the full image slices to fit the new specified dimensions, using no interpolation

and the 'nearest' mode to maintain any displacement discontinuities and minimise any blurring at the edges of the structures when resizing.

The binary mask used by the model is downsized to the output dimensions as it is used within D^2IM to suppress any unneeded output nodes in the predicted displacement fields. To provide consistent visualisation of the results, an additional binary mask, altered by eroding the downsized input mask by 2 pixels, was created to cut off the edge of both strains and errors for visualisation purposes. This serves to exclude unrealistic values of local prediction based on DVC measurement at the borders, which are known to be affected by errors due to progressively reduced amount of greyscale information in those regions [37]. Furthermore, it should mitigate any errors arising due to a mismatch between the input mask and DVC output that may occur due to differences between the single mask slice used as a D^2IM input and the 50 mask slices considered by the DVC correlation window when measuring displacements.

To ensure the ML model is learning sensible relationships, sets of image slices were removed from the data if they contained minimal bone structures that did not span the cranio-caudal dimension of the slice leaving a final dataset size of 251 image sets. Due to the relatively small dataset, the judgement of which slices to remove was done manually based on a visual inspection of the input slices. Finally, the greyscale (pixel values 0–255) unloaded tomography images were normalised so that their values range between 0 and 1.

2.4. Machine Learning Model D^2IM structure, training and strain prediction

The deep learning (DL) model D^2IM is a feed-forward convolutional neural network (CNN) that takes an unloaded tomography and a binary mask as an input to predict the voxel displacement fields $\bar{u}$, $\bar{v}$, and $\bar{w}$. The displacement field slices obtained from DVC were used as the ground truth for training and evaluating the accuracy of the model. It should be noted that while all cases presented in this paper have used the binary mask as an input to improve accuracy of the predictions, the framework can be easily tuned to work without a mask input if required.

The architecture of D^2IM as feed-forward CNN was loosely inspired by VGG16 [38], containing multiple stacked convolutional blocks of increasing depth separated by max pooling layers feeding into multiple dense (fully connected) layers. The key difference comes in the output layer, as rather than being used for the classification of images, D^2IM is used for regression, predicting the values for the three displacement fields based on the image mappings learnt in the convolutional layers.

D²IM was created in Python using the machine learning platform TensorFlow [39] with a network architecture visually summarised in the right-hand portion of Fig. 1 and with additional details in Table 2. Pairs of convolutional layers are stacked using 3×3 convolutions with zero padding and a stride of 1, followed by a max-pooling layer with a 2×2 pooling window with a stride of 2. Four convolutional stacks are used, with the depth of the convolutional layers progressing from the first stack with a kernel size of 32 with each progressive stack doubling in size to achieve kernel size of 256 in stack 4. Once flattened, this is followed by three dense layers of 512 channels, with the third dense layer being fed into the three 20×20 output layers with a multiply function applied to the binary mask that removes any values predicted outside the region of interest. These three outputs predict individual displacement windows that can be reconstructed into the 20×20 fields *u*, *v*, and *w* corresponding to the image provided for the input scan experiencing a 6500 N axial load. This model has 36395282 trainable parameters in total, uses ReLU activation functions for every layer (output layers have no activation as this is a regression problem) and uses batch normalisation before every layer. The dense layers are regularised to minimise overfitting using dropout layers, with a rate of 0.5 and L2 regularisation using a factor (λ) of 0.001.

To train the model, the dataset was shuffled and split into three sets – training (80 %), validation (10 %) and test (10 %), where the training set was used to train the model parameters, the validation set to monitor the model's performance during training and the test dataset is reserved for the final evaluation of the model's accuracy. D²IM was trained for 500 epochs total with a batch size of 100 using the Adam optimiser to minimise the combined Mean Squared Error (MSE) for predictions of the three displacement components. A learning rate schedule was used, such that the first 300 epochs used a learning rate of 0.001 and the remaining 200 used a learning rate of 0.0001; however, the MSE for the validation dataset seemed to plateau shortly after the 400th epoch. The model training took approximately 6 minutes using an NVIDIA RTX A6000 GPU.

Strain fields can be calculated by taking spatial derivatives of the displacements from both the DVC-measured displacements and predicted using D²IM as an alternate way of assessing the performance of the model's predictions of mechanical behaviour. For this study we focused on both quantitative and qualitative comparisons of the normal strains in z (ε_{zz}), which is the axis of largest displacement due to the experimental setup.

2.5. Adjustments for clinical imaging

To explore how well the trained D²IM model could be generalised to

Table 2
D²IM network Architecture, where 'maps' gives the number of output feature maps and 'k' provides the convolutional window size.

Input1 – Tomography (256×256 Image)		
Convolutional Stack 1:	2D Convolution: maps=32, k=3×3	
	2D Convolution: maps=32, k=3×3	
	Max-pooling: 2×2	
Convolutional Stack 2:	2D Convolution: maps=64, k=3×3	
	2D Convolution: maps=64, k=3×3	
	Max-pooling: 2×2	
Convolutional Stack 3:	2D Convolution: maps=128, k=3×3	
	2D Convolution: maps=128, k=3×3	
	Max-pooling: 2×2	
Convolutional Stack 4:	2D Convolution: maps=256, k=3×3	
	2D Convolution: maps=256, k=3×3	
	Max-pooling: 2×2	
Dense Layer: 512		
Dense Layer: 512		
Dense Layer: 512		
Multiply: Input2 – Mask (20×20 image)		
Output 1 – $\overline{u}$ (20×20 image)	Output 2 – $\overline{v}$ (20×20 image)	Output 3 – $\overline{w}$ (20×20 image)

making predictions from images acquired with a lower visual fidelity (i. e. clinical), some adjustments were made to the processes outlined in previous sections. Clinical tomography (AquilionOne, Toshiba, Japan) using facilities at Northern General Hospital in Sheffield, UK, was performed with the following scanning parameters: voltage: 120 kV, current: 250 mA, convolution filter: FC30, achieving a voxel size of 190x190x500μm on the same vertebrae previously examined using high-resolution XTC by Palanca et al. [8,30]. However, these clinical images were only used to capture the vertebrae pre-compression and before the artificial lesions were added, meaning that only intact cases can be examined, and the ground truth DVC measurements obtained from the high-resolution XCT images must be used to evaluate the predictions from the low-resolution input. Data for an additional test was created to add an intermediate state of image fidelity, where a high-resolution tomography was downscaled to match the resolution of the clinical image, creating a case where more of the microstructure was captured, albeit with less detail. The images from the two tomograms (clinical CT and downscaled XCT) were then resized to be 256×256 voxels, as in Section 2.3, to provide a consistent input to D²IM.

Consequently, a new iteration of D²IM was trained with an aim of increasing the ability of the model to generalise to lower resolution inputs. The details of the model architecture and training process are identical to those outlined in Section 2.4 apart from the inclusion of a data augmentation step within the data generator being used during training. This augmentation takes the normalised images and adds a random brightness adjustment (up to ±0.1), contrast shift (up to 0.95–1.05) and a blur (kernel window ranging 1–12 pixels). This serves to make the model less sensitive to the small variations in image brightness that can occur during imaging, with the blurring making the model less sensitive to changes in the microstructure.

2.6. Error Metrics

Multiple indicators were computed to quantify the errors of predictions made using D²IM:

- The performance of the model predictions for each type of displacement field could be individually measured using a Root Mean Squared Error (RMSE):

$$RMSE_u = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(u_i - \overline{u}_i\right)^2}$$

$$RMSE_v = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(v_i - \overline{v}_i\right)^2}$$

$$RMSE_w = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(w_i - \overline{w}_i\right)^2}$$

Where N is the number of windows where a displacement value is being predicted; u_i, v_i and w_i are the measured displacements in that correlation window; and $\overline{u}_i$, $\overline{v}_i$ and $\overline{w}_i$ are the predicted displacements.

- The overall performance of the model predictions is monitored and evaluated using the sum of the Root Mean Squared Errors of all three displacement fields:

$$TRMSE = RMSE_u + RMSE_v + RMSE_w$$

- The performance of the model predictions of each type of displacement and strain field were assessed using a relative error as a percentage of the measured value for each predicted window *i*, with

these error distributions being used to obtain and visualise the mean and standard deviation of the prediction error. For visualisation of the errors in the full fields this is capped at 100 % to prevent outliers from dominating the plots.

$$RE_u^i = \left| \frac{u_i - \overline{u}_i}{u_i} \right|, \quad RE_v^i = \left| \frac{v_i - \overline{v}_i}{v_i} \right|, \quad RE_w^i = \left| \frac{w_i - \overline{w}_i}{w_i} \right|$$

- A Mann Whitney U test with a significance level $\alpha = 0.05$ was used to compare the variation of the distributions of displacement fields measured and predicted. A non-parametric test was chosen as there was no indication that the samples being compared were normally distributed. A two-sided version of the test was used to compare the measured displacements using different DVC methods and for the predicted vs measured displacements. A one-tailed test was used to compare the relative error distributions for different displacement components.
- The correlation coefficient (R^2) between the measured and predicted displacements were calculated for each displacement component as an alternate means of assessing overall performance in predicting each component.

3. Results and Discussion

While both the DVC measurements and D^{2}IM predictions represent the displacement in voxels, the results herein presented have been converted to their true displacement values in μm, using the conversion rate of 1 voxel = 39μm as achieved in the original experiment by Palanca et al. [8,30].

The ability of D^{2}IM to generalise across the three displacement components was first examined by making a comparison of the RMSE that has been summarised in Table 3. Here it can be observed how both validation and test predictions do have a notably higher error than the training data, but the training curves indicated that there was no systematic overfitting of the validation data, which only ever improved before it plateaued. The most significant contribution to this increase in overall RMSE comes from the v displacement predictions, with validation and test errors many times larger compared to the training results. The validation and test RMSE for u displacement predictions were also larger than those observed for the training set, but their contribution to the overall error was comparatively minor. The w displacements demonstrated the most consistent test prediction behaviour by this metric as while the validation error was nearly double that of the training data, the RMSE for the test dataset is just over one fifth of a voxel larger that the validation set error, which is the smallest increase in error between the validation and tests sets across the three displacement components, being easily explainable by variation in the samples and seeming to indicate that there is no systematic overfitting for this displacement component.

However, the raw RMSE overlooks the fact that components with measured values of a higher magnitude would be expected to have an intrinsically higher RMSE regardless of the relative precision of the predictions. Hence, summary statistics describing the distribution of measured and predicted displacements in the test dataset (excluding predictions of zero in regions outside of the mask) are summarised in Table 4.

The measured and predicted displacements for all components were found to be significantly different across the entire population of the test dataset (p<0.05 using a Mann Whitney U test), as predictions failing to reach the more extreme values observed in the measurements often led to significant differences in the mean value and standard deviation calculated, with the median generally providing a closer match (except for v). Regardless, these statistics seem to demonstrate the fundamental mechanical behaviour is being captured for all three displacement components, with differences between the measured and predicted mean/median values measured ranging 4–35μm (with a 5–22 % relative error other than the 82 % relative error for the median of v). At the same time, it does also reinforce the overall trends observed in the RMSE, with the $\overline{u}$ displacement's relatively small RMSE being partially explained with the comparatively small mean value and tighter standard deviation. The $\overline{v}$ displacements have by far the largest range in measured values, which could explain why this component had the largest observed RMSE even for the training predictions (Table 3). The poor behaviour in the test predictions for $\overline{v}$ can also be observed with the statistics having some of the biggest discrepancies between the measured and predicted values, with a standard deviation over double that of $\overline{w}$, despite being broadly comparable in mean displacement. Finally, the $\overline{w}$ displacements continued to show more consistent prediction behaviour with the standard deviation being closest, in both relative and absolute terms, to the corresponding measured value. Some clear discrepancies can be seen between the predicted and measured minimum and maximum values for w, which is consistent with a typical trend of under-predicting high values and over-predicting low values already reported in literature [26,40], which skews the mean prediction leading to a 19 % error, but when mitigating the influence of these outlier values by considering the median a more modest 5 % error is found.

A correlation analysis (Fig. 3) was performed between all the measured and predicted displacements to further explore the error in the predictions. This again highlighted that the predictions were generally accurate, with strong positive correlation coefficients for all three displacement components. While there are clear outliers and scatterings of poor predictions, they mostly make up a relatively minor proportion of the 5797 non-masked displacement windows being predicted and plotted. The $\overline{u}$ displacements had the lowest R^2 value of 0.82, with the most egregious scattering of poor predictions, especially for the more extreme valued displacements. The prediction correlations for the other displacement components are more comparable, with R^2 value of 0.91 for $\overline{v}$ and 0.94 for $\overline{w}$.

Inspecting the scatter charts, the w displacements seemed to produce the most consistent behaviour overall, as while very tight in some places, the $\overline{v}$ displacements displayed some clear clustering of poor predictions, along with more significant outliers.

The points on this scatter chart have been coloured according to the vertebra sample they originated from. It should be noted that, due to random sampling, different numbers of slices from each case are present in the test data with no example slices from the second intact specimen. This serves to highlight that many of the notable regions of incorrect predictions arise from the same samples, notably often from cases where more extreme displacements were measured. The second anterior lesion (orange) and first lateral lesion (grey) cases both have some of the highest displacement magnitudes, but also a very large variation in prediction accuracy compared to most other cases, likely due to the lower number of lesioned examples in the data. The intact cases generally register both smaller displacement and more accurate predictions except for intact case one (red), with large displacements and scattering in u and w, but near zero displacements being accurately predicted in v. This may be due to a different mechanical pattern compared to the other cases where the load transfer was primarily in compression, leading to some uncertainties in the model prediction as is attempts to learn both this rare behaviour and the more dominant patterns seen in the other intact cases.

Table 3
RMSE for D^{2}IM voxel displacement predictions for the training, validation, and test data after training for 500 epochs.

RMSE (voxels)	Train	Validation	Test
$\overline{u}$	0.5285	0.8514	1.4342
$\overline{v}$	1.5176	3.7724	4.9117
$\overline{w}$	1.0111	1.771	1.9947
Overall	3.0572	6.3948	8.3406

Table 4
Summary statistics for the measured and predicted displacement fields of the test dataset.

	Measured u (µm)	Predicted $\bar{u}$ (µm)	Measured v (µm)	Predicted $\bar{v}$ (µm)	Measured w (µm)	Predicted $\bar{w}$ (µm)
Mean	74.05	57.85	156.69	166.18	-184.82	-149.21
Median	39.91	35.11	-29.42	-5.00	-159.12	-142.61
St. Dev.	111.15	66.47	508.51	399.04	201.86	165.82
Min	-284.65	-113.18	-999.18	-858.67	-801.79	-640.35
Max	509.95	331.35	1450.48	1262.42	555.21	458.08
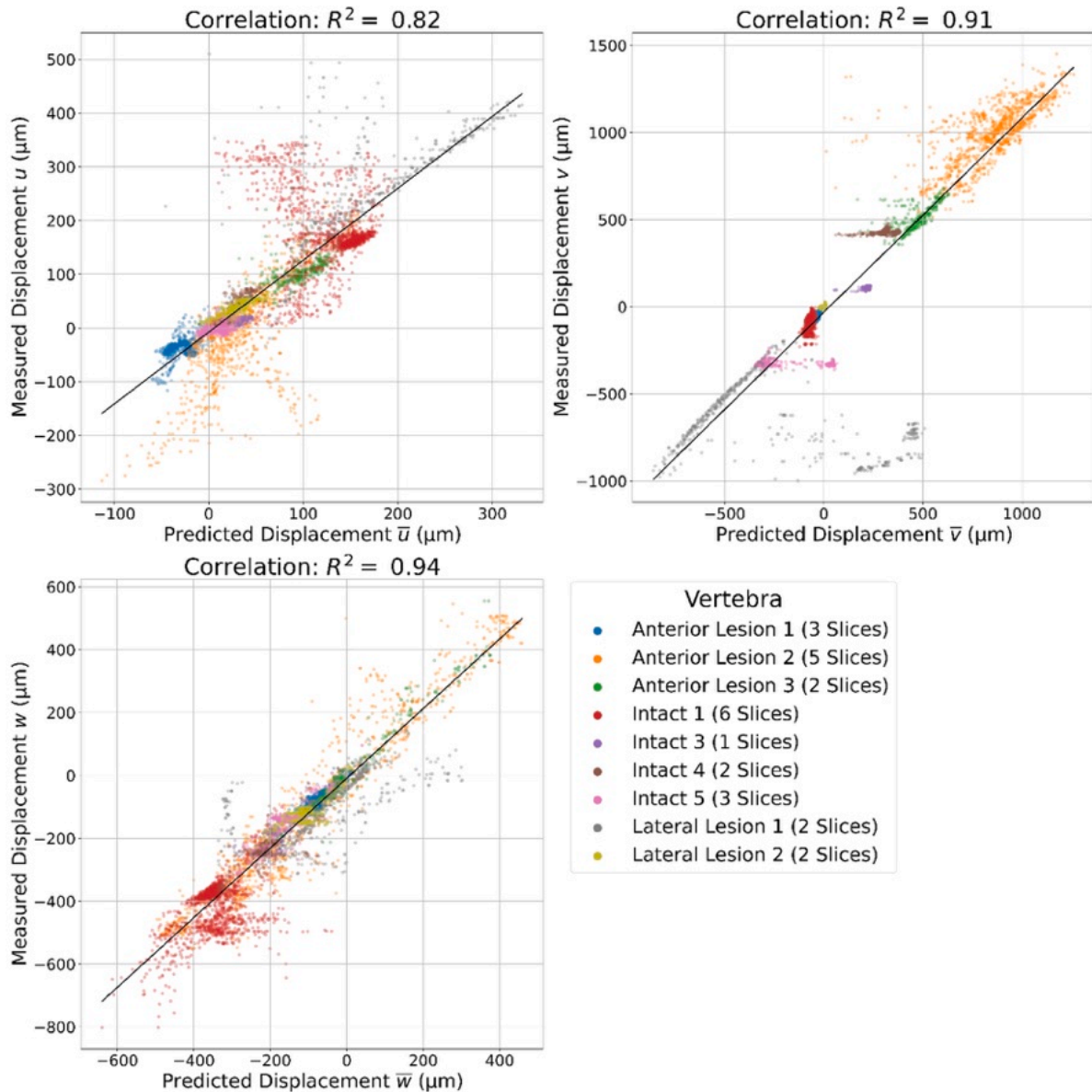
Fig. 3. Correlation analysis between all measured and predicted displacements for the test dataset, performed across all three displacement components, including a line of best fit and the calculated correlation coefficient R^2. Points have been coloured according to the vertebra they belong to.

The relative error of the predictions was also examined and reported in Fig. 4, where the error of a predicted displacement is represented as a percentage of the measured value.

These results reinforced previous findings as most predictions for all three displacements were fairly accurate, with the distribution of relative displacement errors having median values clearly skewed towards zero (31 % for $\bar{u}$, 17 % for $\bar{v}$ and 15 % for $\bar{w}$) and interquartile ranges of 63 % for $\bar{u}$, 32 % for $\bar{v}$ and 26 % for $\bar{w}$. Meaning that once again $\bar{w}$ continued to show the most consistent behaviour and the tightest error

distribution, with the range of the recorded errors being significantly smaller than those observed for the other two displacement components ($p \ll 0.05$ using a Mann Whitney U test). While all displacement components predicted by D^{2}IM seemed to capture the overall relationships with a reasonable degree of accuracy, these metrics consistently showed the $\bar{w}$ displacements as the best performing component. This is both encouraging and expected considering the nature of the problem, as the framework implemented in D^{2}IM for this study is examining cases of a bone loaded in compression, where the largest displacements are
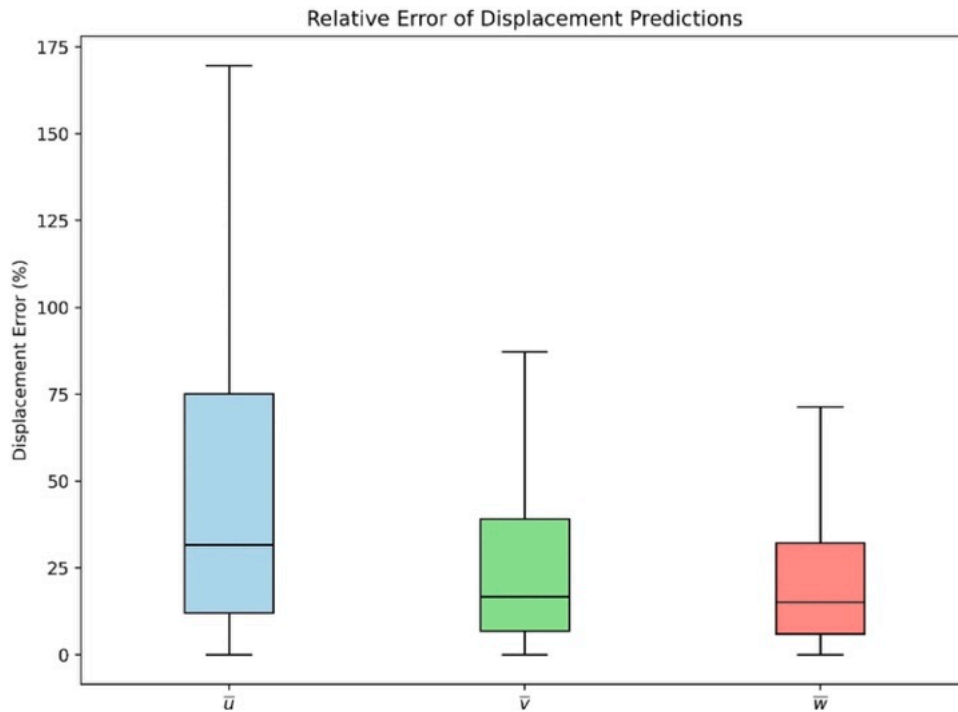
Fig. 4. Box and whisker chart with relative error distribution as a percentage of the measured value across the three displacement components. Boxes show the interquartile range and whiskers the maximum and minimum observed errors (excluding outliers).

expected in z and normal strain fields in that direction (ε_{zz}) are used in describing the local mechanics [41,42].

Plots of all displacement components and their errors for all images in the test dataset can be found in the supplementary material. However, as an in-depth analysis of all components for the entire test dataset is impractical within the scope of this paper, indicative results for four different cases present in the dataset have been chosen and reported in Figs. 5–8, where the ability of D^2IM to predict full-field displacements ($\overline{w}$) and strains ($\overline{\varepsilon}_{zz}$) is shown for intact and lesioned vertebral sections sliced in both anterior-posterior and left-right directions.

The first image is of an intact anterior-posterior sliced vertebra as shown in Fig. 5. D^2IM is predicting very similar behaviour for both displacement and strain fields to those measured using DVC. The displacements both showed the largest magnitudes ($\sim$366μm for w and $\sim$338μm for $\overline{w}$) at the top endplate and a relatively consistent displacement throughout most of the vertebral body, before rapidly shrinking to nearly zero towards the bottom endplate. This was matched by both strain profiles having a bar of similar high magnitude strain at the top and bottom endplates ($\sim$48532μe for ε_{zz} and $\sim$33660μe for $\overline{\varepsilon}_{zz}$), with comparatively small strain being registered throughout the rest of
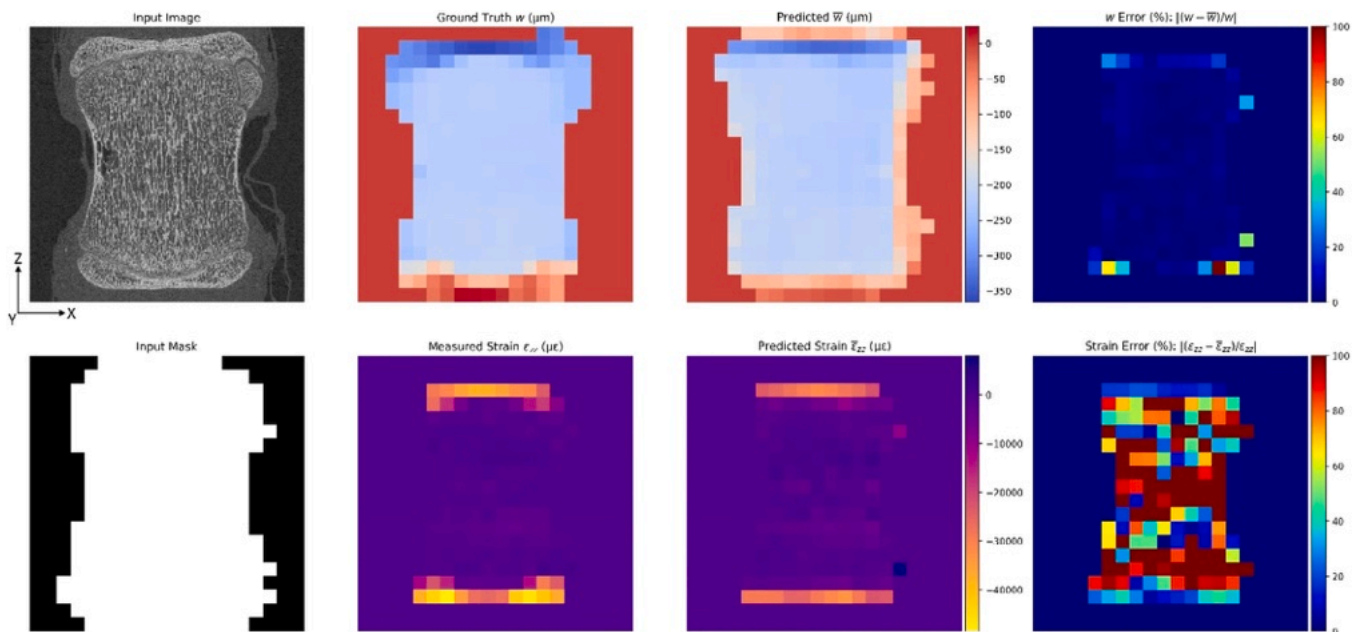

Fig. 5. Results for an intact anterior-posterior sliced tomography. Top row left-right: Input tomography slice, measured displacements, predicted displacements and relative error of displacement. Bottom row left-right: Input binary mask slice, measured strain, predicted strain and relative error of strain.
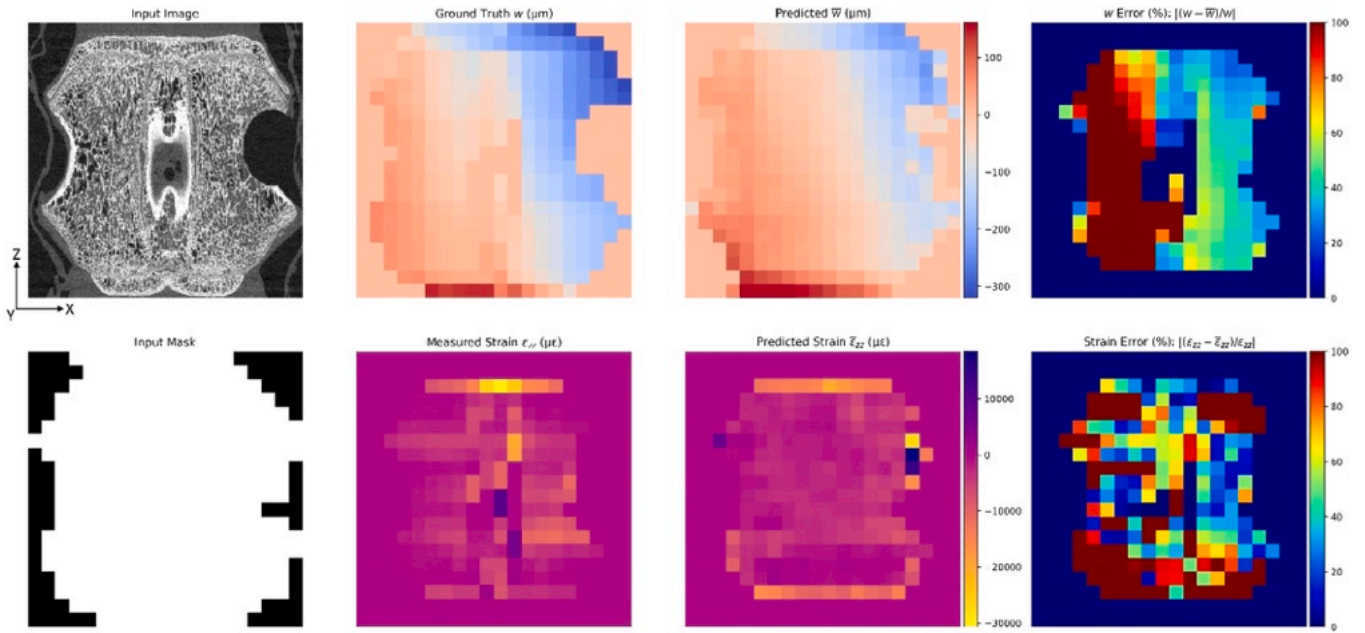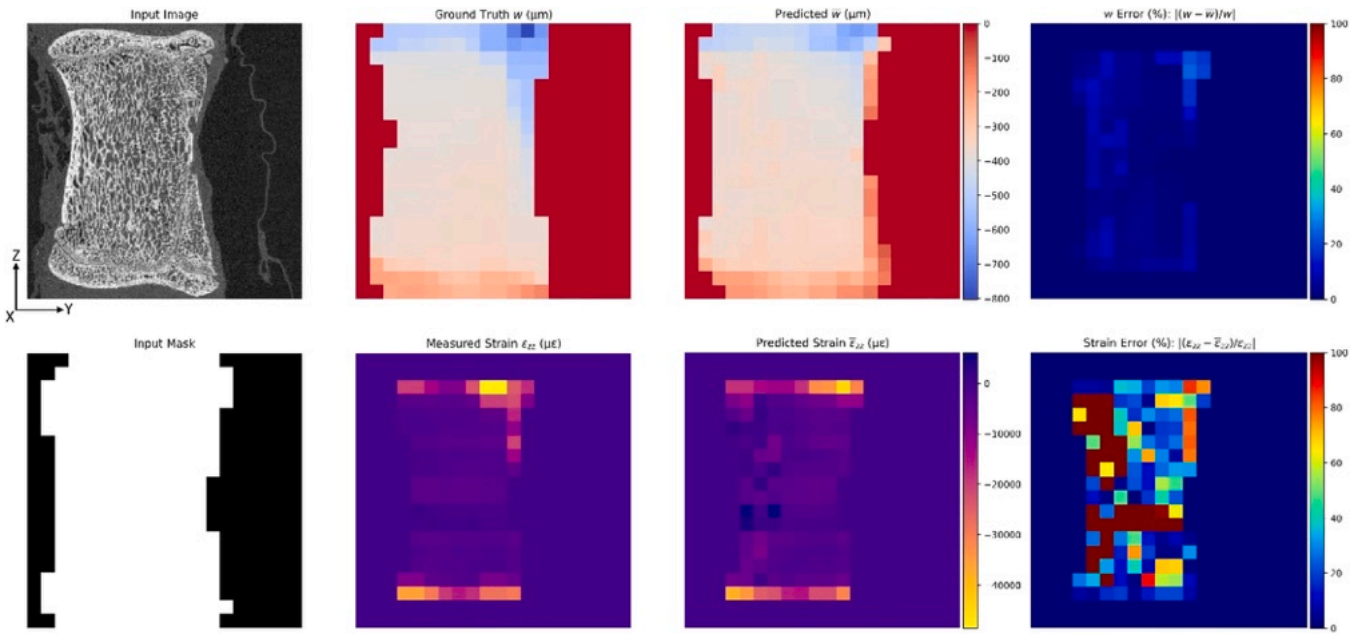
Fig. 6. Results for an anterior-posterior sliced tomography with lesion. Top row left-right: Input tomography slice, measured displacements, predicted displacements and relative error of displacement. Bottom row left-right: Input binary mask slice, measured strain, predicted strain and relative error of strain.

Fig. 7. Results for a left-right sliced tomography with no lesion. Top row left-right: Input tomography slice, measured displacements, predicted displacements and relative error of displacement. Bottom row left-right: Input binary mask slice, measured strain, predicted strain and relative error of strain.

the structure. This mechanical behaviour is likely due to the growth plates made of softer tissues that can be found at the ends of the vertebra. Displacement errors were generally small and without much noticeable pattern within the vertebral body with a median error of 1.6 % (81 % for strain), except at the bottom of the structure where there is a slight mismatch in the displacement and strain behaviour, being more complicated than a single bar in the measured results. The relative errors in the strain initially seemed large, but for most of the structure this is noise as strains comparatively small, while in regions of large strain around the two end plates the relative error is much less significant. Some outliers were localised at the borders/corner of the masked image and, despite the strategy of using an additional mask to

remove boundary errors, there may be instances in the visualisation with errors still caused by either DVC boundary uncertainties in the measurement or mismatch occurring between the masks.

The second case shown in Fig. 6 is on another anterior-posterior sliced image from a vertebra with artificial lesion. D^{2}IM captured the underlying displacement and strain behaviour, with the lesioned side of the structure experiencing negative displacements in z, while those on the non-lesioned side were small, with some positive values. The strain was quite small for both the measured and predicted output, barring a line of high strain magnitude ($\sim$30661$\mu\varepsilon$ for ε_{zz} and $\sim$19932$\mu\varepsilon$ for $\overline{\varepsilon}_{zz}$) at the top endplate and some outliers visibly caused in this case by a masking mismatch. While a visual inspection of the fields indicates that
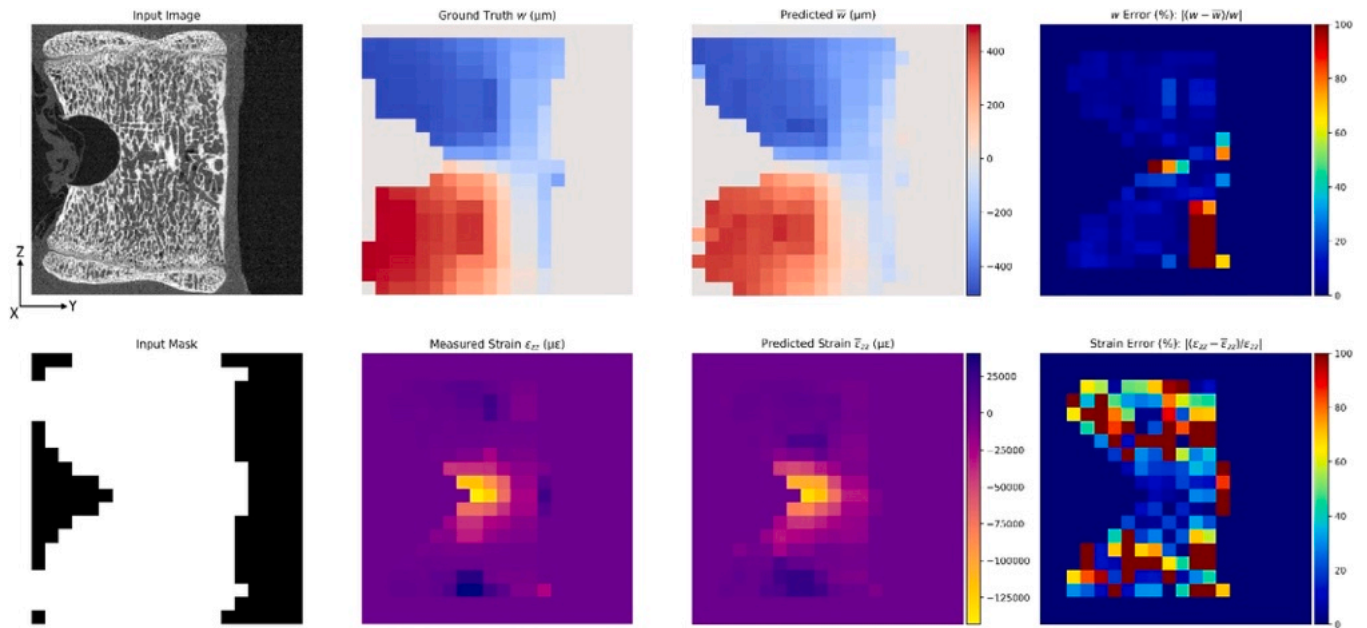
Fig. 8. Results for a left-right sliced tomography with lesion. Top row left-right: Input tomography slice, measured displacements, predicted displacements and relative error of displacement. Bottom row left-right: Input binary mask slice, measured strain, predicted strain and relative error of strain.

the underlying mechanical behaviour is being captured, there was more of a systematic error when compared to the intact case in Fig. 5, with a significantly higher median error of 70.4 % for displacement and 71.2 % for strain, though it should be noted that the region of more than 100 % relative error on the left side of the vertebra registered near zero displacements, so the absolute error in this region is comparatively small. This is believed primarily to be an issue of numerosity, as while both five cases with and without lesions were used, the position of the lesion in the vertebra differed between the lesioned cases. Specifically, three of the lesions were in the anterior position of the vertebra and only two had a lesion in a lateral position as is presented here. Consequently, due to having three less directly applicable cases to learn from, when compared to an intact case, it is unsurprising that D²IM found the lateral lesion scenario to be more challenging to accurately predict.

In Fig. 7 an image of an intact vertebra sliced in the left-right direction is presented. Similarly to the anterior-posterior case, there was a very good agreement between prediction and measurement. Both identified a band of large negative displacement at the top endplate of the vertebra (maximum ~802μm for w and ~640μm for $\overline{w}$), with a point of high displacement at the top-right. Most of the structure showed relatively consistent smaller displacements, before reaching a band of nearly zero displacements along the bottom endplate of the vertebra. It must be noted that while D²IM identified the top right being the point of highest displacement magnitude with a band of large displacements going down the right edge, the predictions were smaller in magnitude than were measured. This error was carried across into the strains, where both the predicted and measured strain fields showed similar bands of strain at the top and bottom endplate, but the predictions did not capture the vertical band of strain going down the top-right edge of the vertebra. Despite some larger errors in this region, with median errors of 2.7 % for displacement and 33 % for strain (strain error again somewhat misleading as it mostly originates from the region of near zero strain), the fundamental mechanical behaviour is clearly being captured.

The final case being presented (Fig. 8) is on the prediction results for a vertebra with a lesion sliced in a left-right direction. The results showed very good agreement in the general behaviour, with less of a notable systematic error than in the other case presented with a lesion in Fig. 6, likely due to the presence of an additional case with an anterior

lesion in the dataset. While the displacements were generally slightly smaller in magnitude for the predictions, the underlying behaviour is clearly being caught with negative displacements above the lesion and positive displacements below, both increasing in magnitude as they move towards the front of the vertebra. The displacement errors are generally small and seemingly random, albeit larger overall when compared to the intact cases (median error of 5.8 %), likely related to the dataset including less cases with an anterior lesion than of intact vertebra. The strains reported a similar good match between the predicted and measured values, with both cases showing a large negative strain near the lesion in the centre of the vertebra (~143179με for ε_{zz} and ~131089με for $\overline{\varepsilon}_{zz}$), with some smaller, but still significant positive strains near the top and bottom endplates. The strain relative error has a median value of 38 %, however it again shows low errors in the regions where high strains are measured and no noticeable error pattern elsewhere.

While only presenting the ε_{zz} strain output, the local mechanics predicted by D²IM is in line with that presented other studies using DVC on intact vertebrae [30,43], where high compressive strain localised in correspondence to endplates and reaching magnitudes up to 55000με. In the presence of lesions. Palanca et al. reported high localised strains, with magnitudes in the range 80000–200000με [8], which is consistent with D²IM prediction. A further consideration is needed for predicted strain fields. Given the strategy used to increase numerosity with 2D XCT slices and the conscious choice of structuring D²IM to use the primary DVC output, the normal strains in w are being derived directly by taking spatial derivatives of the predicted displacement fields and are not themselves 'learned', which similar DL studies have found to be a less accurate way of predicting strain [17]. Thus, they carry across the displacement errors, mainly resulting in some underestimation of higher strain magnitudes and some overestimation of lower strain magnitudes. Whilst the former generally remains representative of regions of tissue already exceeding yielding (i.e. 10000με for bone in compression [44]), the latter could indicate a prediction of yielded portions of the tissue that weren't measured as such. However, despite the presence of few areas experiencing high percentage of relative error between measured and predicted strain, the mechanical significance of unyielded bone tissue was broadly met (Fig. 9, top row). Incongruences in strain predictions exceeding the set threshold of 10000με, when compared to measured
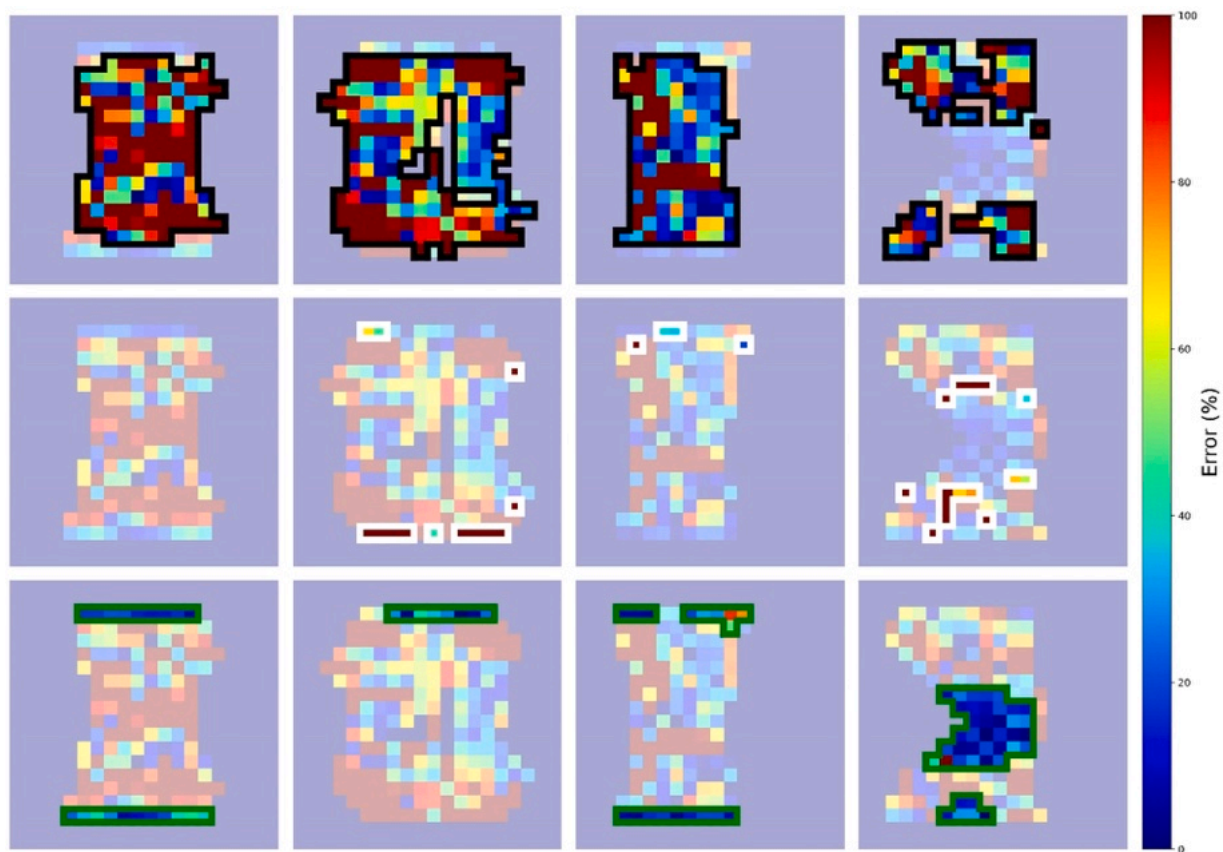
Fig. 9. Relative error of strain predictions for the different cases reported in Figs. 5–8 (left-right). Images have been marked with windows highlighting regions of interest regarding bone yielding behaviour. Top: marked in black is measured strain in DVC windows lower than 10000με and predicted strain also lower than 10000με. Middle: marked in white is measured strain in DVC windows lower than 10000με but predicted strain higher than 10000με. Bottom: marked in green is measured strain in DVC windows higher than 10000με and predicted strain also higher than 10000με.

values which did not, were mostly localised in border outliers and in proximity of regions where bone was either yielded or pre-yielding (Fig. 9, middle row). Much lower relative errors were found when both measured and predicted strain exceeded 10000με (Fig. 9, bottom row), which overall indicates the ability of D^{2}IM predictions to successfully represent the mechanical state of bone.

While the efficacy of D^{2}IM for predicting displacement fields for the high-resolution input images has been demonstrated, a further key consideration is how well this framework can generalise to making predictions for images acquired with a lower visual fidelity. This would be a key concern for the future implementation of this technology in a clinical setting where the acquisition of images of patient tissue with such detail is currently unattainable.

Fig. 10 reports results for low resolution inputs corresponding to the intact anterior-posterior sliced case, previously examined using high resolution inputs in Fig. 5. While both the downscaled and clinical inputs clearly contain more error in their predictions, the displacements still captured the underlying pattern of high negative displacements at the top of the structure, smaller negative displacements throughout the main body of the vertebra before reaching near zero displacement at the base of the structure. Consequently, while the lack of detail in the inputs has increased the noise in the prediction, it has clearly still been able to capture the fundamental mechanical behaviour.

Fig. 11 shows results for low resolution inputs corresponding to the intact left-right sliced case, previously examined using high resolution inputs in Fig. 7. When comparing the displacement distribution with the high-resolution results the model at first glance appears to have performed well in capturing the displacement distribution within the structure using both the clinical and downscaled input images, with both

having the highest displacements at the top right of the structure, a corresponding high band of displacement across the rest of the top and smaller, but still significant, displacements throughout the rest of the body of the structure. However, when looking at the error there is clearly a systematic underprediction of hundreds of microns throughout the structure leading to errors of 40–70 % throughout the structure.

While there is still an obvious requirement of further development of D^{2}IM to be able to make accurate predictions from lower resolution inputs, it is reassuring that with only a minor amendment of adding data augmentation the model can predict a sensible displacement field that captured the expected fundamental mechanics from significantly poorer quality input data. This provides confidence that, after further experimentation with hyperparameters and model architecture along with the capability to work with full 3D images, D^{2}IM will be capable of making effective predictions from clinical inputs, making it a realistic prospect that this technology could ultimately be deployed in this context.

The D^{2}IM model presented in this paper is the outcome of an initial investigation to ascertain the viability of this framework, and as such limiting assumptions have been made to simplify some tasks.

The simplification producing the biggest impact was the decision to use 2D slices as the inputs to the model. When taking a single slice there are 49 other slices in that DVC correlation window that were used to measure the displacement and not being used by D^{2}IM, meaning a great deal of greyscale information is not considered in the prediction. For example, in this dataset half of the data was from vertebrae with an artificial lesion, which predictably impacted on the observed structural behaviour but, as evidence of this lesion is generally not found in all slices, D^{2}IM may be presented with two almost identical slices of which one was from a lesioned vertebra and the other was not. While this could
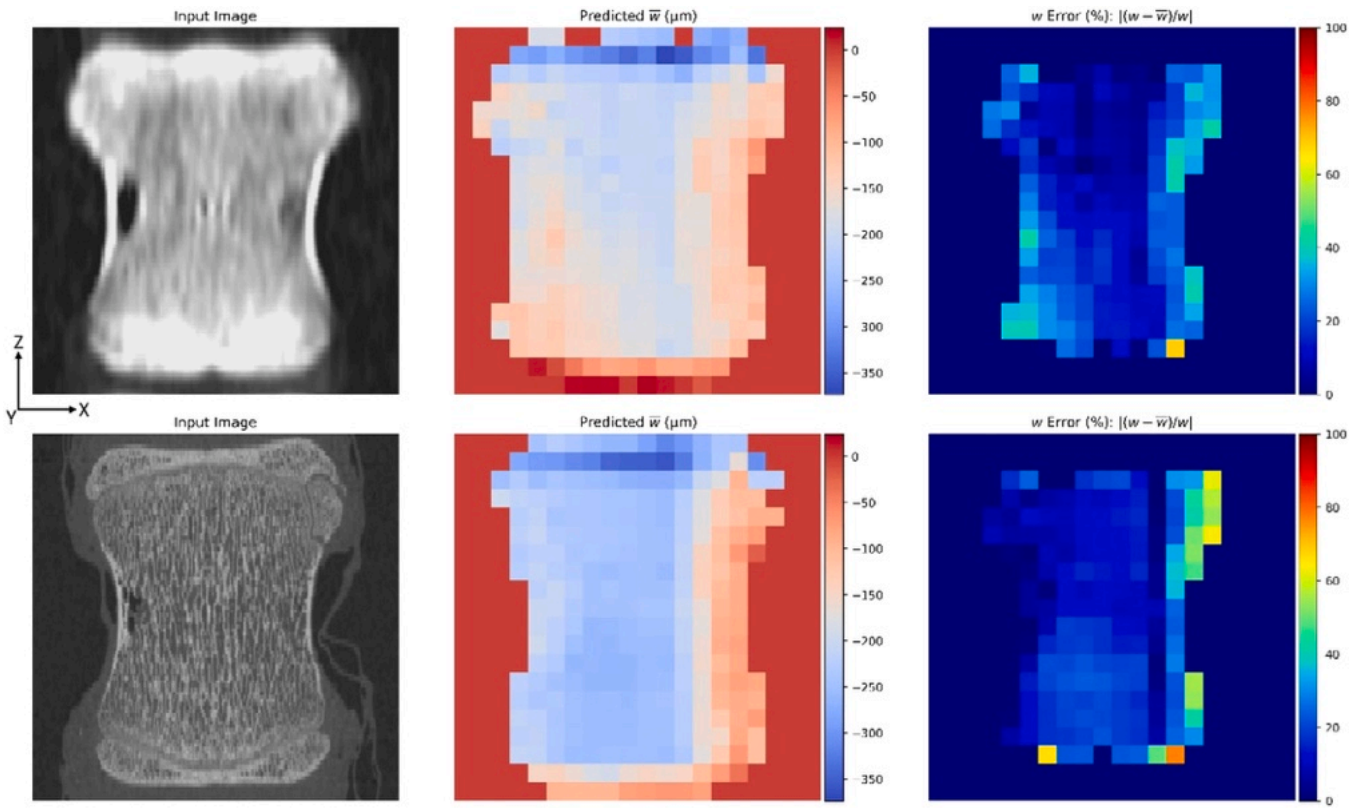
Fig. 10. Results for an intact anterior-posterior sliced tomography. Top row: results for clinical tomography slice. Bottom row: results for downscaled high resolution tomography slice. Left-right: Input tomography slice, predicted displacements and relative error of displacement.
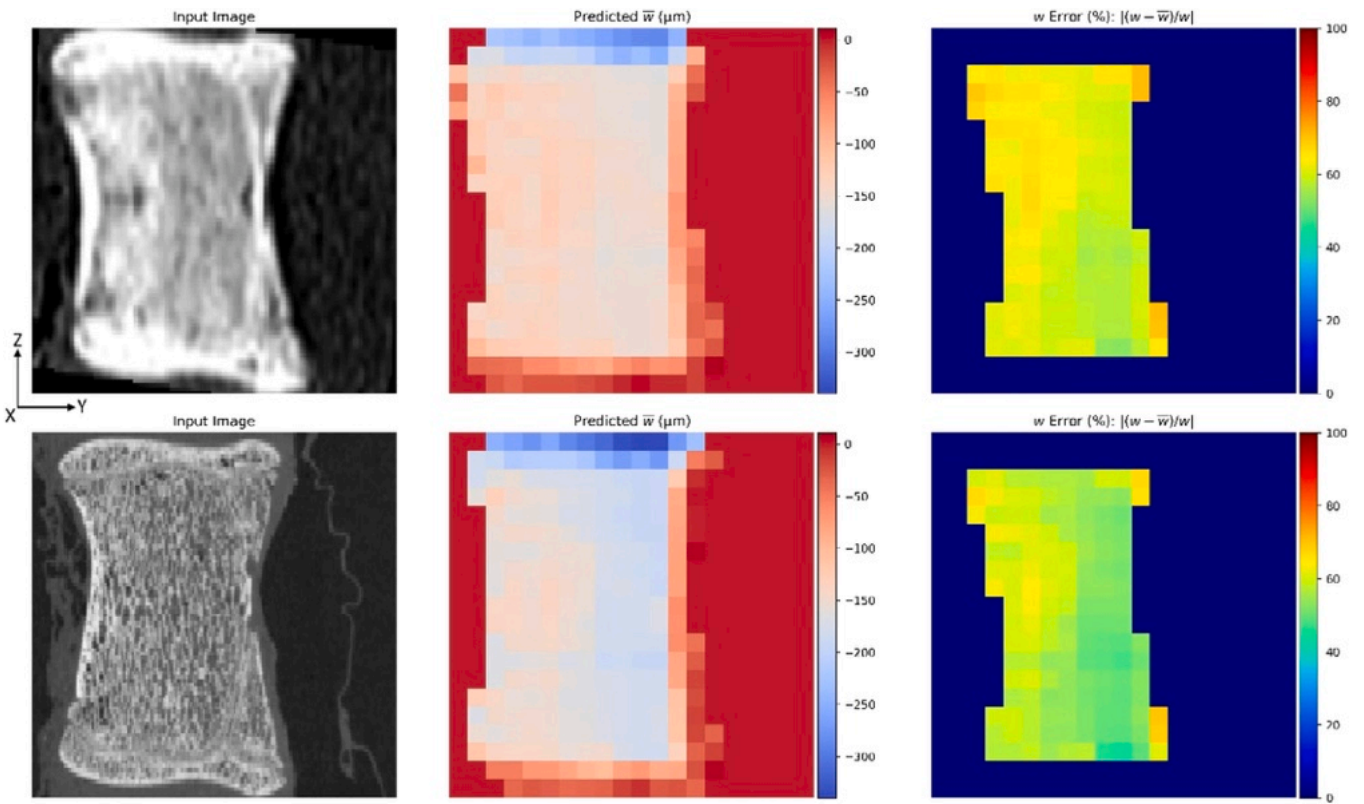

Fig. 11. Results for a left-right sliced tomography with no lesion. Top row: results for clinical tomography slice. Bottom row: results for downscaled high resolution tomography slice. Left-right: Input tomography slice, predicted displacements and relative error of displacement.

represent a potential problem for all displacement components, the u and v displacements were most affected. The w displacements experienced this to a less significant degree as all images in the dataset included the cranio-caudal (z) change in structure, whichever direction slicing was performed. A clear example of this can be found in the low-resolution results shown in Fig. 11, where training data include many left-right slices that look superficially very similar to this case, but this slice comes from the first intact case (red dots in Fig. 3), which registered significantly higher displacements than all other intact cases in the dataset. With the lack of image detail, D^2IM made a more conservative prediction of smaller displacements also reinforced by the mean squared error loss function with an L2 norm, which trained the model to require a higher level of confidence before predicting larger displacement values. This slice simplification also added a limitation to the possible strain analysis that could be performed. Due to the combination of anterior-posterior and left-right slices being used in the dataset, ε_{zz} was the only meaningful strain component to calculate and report in the current setup.

Another limitation was the data splitting strategy utilised, as slices from all scans were randomly shuffled into the training, validation and test datasets; meaning that the test set was not truly independent with a potential of some data leakage arising from image slices in all sets coming from the same original scan. A more robust splitting strategy would have been to hold back an entire set of slices for one or more vertebrae to use as the test data, ensuring the test data to be fully independent. While it is hoped this strategy can be employed in future studies with D^2IM, it was deemed impractical for this proof-of-concept work as the dataset used was already quite small. To fully test the three different loading scenarios (intact, anterior lesion and lateral lesion), a case of each would have to be reserved for the test dataset, leaving only 7 (3 with lesions) cases remaining for training and validation, which was considered insufficient to train a model capable of making sensible predictions for all cases.

The requirement of resizing the image to provide a uniform input and output size for the data also presents a potential limitation, as this resizing transformation will always cause some data loss in how voxels are interpolated. This has been considered a separate issue to the central questions being examined in this paper, and the error added by this process when compared to the displacements registered in the original dimensions has not been investigated, but it is acknowledged that a process for returning displacements to the original dimensions with the minimum error will be a key requirement for any practical implementation of this model.

An intrinsic limitation is that D^2IM can only predict mechanical behaviour it has been trained on. Consequently, while D^2IM can make a prediction using any greyscale input image in theory, due to training data, it will currently try to interpret any greyscale image as vertebral bone structure and will try to predict how it deforms under an axial load of 6500 N. In future D^2IM will be trained with more datasets containing different materials and loading conditions that can be considered when making a prediction. However, as practical experiments can often be quite limited in the amount of data produced by the standards of deep learning datasets, data generation via more traditional numerical modelling such as the finite element method [45] has the potential to supplement the data used by D^2IM to both increasing numerosity for experimentally measured cases and allowing for the inclusion of loading scenarios that haven't yet been experimentally tested.

4. Conclusion

In this study a novel data-driven image mechanics (D^2IM) approach was developed to predict displacement and strain fields directly from undeformed XCT images. D^2IM was trained using experimentally measured displacement fields for vertebrae, with and without artificial lesions, obtained via digital volume correlation (DVC). D^2IM was able to make generally sensible predictions for all displacement fields and

displayed an overall good performance capturing the mechanical state of bone for both displacement and strain fields in the main loading direction of the vertebra. Tests using lower resolution and clinical images captured the fundamental mechanical behaviour and demonstrated the potential use of this approach to leverage measurements made using high quality data to make effective predictions on the low quality inputs found in a clinical setting. The findings represent a breakthrough in the prediction of physical fields by exploiting only the rich information contained in undeformed XCT images. The future development of D^2IM will aim at learning a wider range of mechanical data from both hard and soft tissue in healthy and pathological conditions for accurate prediction of physical fields, avoiding long, repetitive and tissue-damaging experimental campaigns but also to advance integration of data in more sophisticated FE simulations. Ultimately, translating into clinical imaging to develop the next generation of diagnostic techniques.

CRediT authorship contribution statement

Peter Soar: Writing – review & editing, Writing – original draft, Visualization, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Marco Palanca:** Writing – review & editing, Investigation, Data curation. **Enrico Dall'Ara:** Writing – review & editing, Investigation, Data curation. **Gianluca Tozzi:** Writing – review & editing, Writing – original draft, Methodology, Formal analysis, Conceptualization.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.eml.2024.102202.

References

[1] G. Tozzi, M.J. Buehler, Advanced Mechanics of Hard Tissue Using Imaging-Based Measurements and Artificial Intelligence. Comprehensive Mechanics of Materials, Elsevier MRW, 2024.
[2] E. Dall'Ara, G. Tozzi, Digital volume correlation for the characterization of musculoskeletal tissues: Current challenges and future developments, Front. Bioeng. Biotechnol. 10 (2022) 1010056, https://doi.org/10.3389/fbioe.2022.1010056.
[3] M. Pena Fernandez, A.P. Kao, R. Bonithon, D. Howells, A.J. Bodey, K. Wanelik, et al., Time-resolved in situ synchrotron-microCT: 4D deformation of bone and bone analogues using digital volume correlation, Acta Biomater. 131 (2021) 424–439, https://doi.org/10.1016/j.actbio.2021.06.014.
[4] A. Karali, E. Dall'Ara, J. Zekonyte, A.P. Kao, G. Blunn, G. Tozzi, Effect of radiation-induced damage of trabecular bone tissue evaluated using indentation and digital volume correlation, J. Mech. Behav. Biomed. Mater. (2023) https://doi.org/10.1016/j.jmbbm.2022.105636.
[5] G. Tozzi, M. Pena Fernandez, S. Davis, A. Karali, A.P. Kao, G. Blunn, Full-field strain uncertainties and residuals at the cartilage-bone interface in unstained tissues using propagation-based phase-contrast XCT and digital volume correlation, Mater. (Basel 13 (2020) 2579, https://doi.org/10.3390/ma13112579.
[6] E. Dall'Ara, A.J. Bodey, H. Isaksson, G. Tozzi, A practical guide for in situ mechanical testing of musculoskeletal tissues using synchrotron tomography, J. Mech. Behav. Biomed. Mat. 133 (2022) 105297, https://doi.org/10.1016/j.jmbbm.2022.105297.

[7] M. Pena Fernandez, C. Black, J. Dawson, D. Gibbs, J. Kanczler, R. Oreffo, G. Tozzi, Exploratory full-field strain analysis of regenerated bone tissue from osteoinductive biomaterials, Materials 13 (1) (2020) 168.

[8] M. Palanca, S. Oliviero, E. Dall'Ara, MicroFE models of porcine vertebrae with induced bone focal lesions: Validation of predicted displacements with digital volume correlation, J. Mech. Behav. Biomed. Mat. 125 (2022) 104872, https://doi.org/10.1016/j.jmbbm.2021.104872.

[9] M.C. Costa, G. Tozzi, L. Cristofolini, V. Danesi, M. Viceconti, E. Dall'Ara, Micro Finite Element models of the vertebral body: Validation of local displacement predictions, PLoS One 12 (2017) e0180151, https://doi.org/10.1371/journal.pone.0180151.

[10] X. Duan, J. Huang, Deep learning-based digital volume correlation, Extrem. Mech. Lett. 53 (2022) 101710, https://doi.org/10.1016/j.eml.2022.101710.

[11] P.S.Q. Yeoh, K.W. Lai, S.L. Goh, K. Hasikin, Y.C. Hum, Y.K. Tee, S. Dhanalakshmi, Emergence of Deep Learning in Knee Osteoarthritis Diagnosis, Comput. Intell. Neurosci. 2021 (2021) 4931437, https://doi.org/10.1155/2021/4931437. PMID: 34804143; PMCID: PMC8598325.

[12] S. Chin-yun Shen, M. Pena Fernandez, G. Tozzi, M. Buehler, Systematic exploration of applied machine learning for bone mechanics, JMBBM (2021), https://doi.org/10.1016/j.jmbbm.2021.104761.

[13] F. Buccino, I. Aiazzi, A. Casto, B. Liu, M.C. Sbarra, G. Ziarelli, G. Banfi, L. M. Vergani, The synergy of synchrotron imaging and convolutional neural networks towards the detection of human micro-scale bone architecture and damage, J. Mech. Behav. Biomed. Mater. (2023), https://doi.org/10.1016/j.jmbbm.2022.105576.

[14] R. Caron, I. Londono, L. Seoud, I. Villemure, Segmentation of trabecular bone microdamage in Xray microCT images using a two-step deep learning method, J. Mech. Behav. Biomed. Mater. (2023), https://doi.org/10.1016/j.jmbbm.2022.105540.

[15] S. Boukhtache, K. Abdelouahab, F. Berry, B. Blaysat, M. Grédiac, F. Sur, When deep learning meets digital image correlation, Opt. Lasers Eng. 136 (2021) 106308, https://doi.org/10.1016/j.optlaseng.2020.106308.

[16] X. Duan, H. Xu, R. Dong, F. Lin, J. Huang, Digital image correlation based on convolutional neural networks, Opt. Lasers Eng. (2023), https://doi.org/10.1016/j.optlaseng.2022.107234.

[17] R. Yang, Y. Li, D. Zeng, P. Guo, Deep DIC: Deep learning-based digital image correlation for end-to-end displacement and strain measurement, J. Mater. Process. Technol. (2022), https://doi.org/10.1016/j.jmatprotec.2021.117474.

[18] W. Muhammad, A.P. Brahme, O. Ibragimova, J. Kang, K. Inal, A machine learning framework to predict local strain distribution and the evolution of plastic anisotropy & fracture in additively manufactured alloys, Int. J. Plast. 136 (2021 Jan 1) 102867.

[19] Q. Gao, H. Lin, H. Tu, H. Zhu, R. Wei, G. Zhang, X. Shao, A robust single-pixel particle image velocimetry based on fully convolutional networks with cross-correlation embedded, Phys. Fluids 33 (12) (2021) 127125, https://doi.org/10.1063/5.0077146.

[20] Y.C. Hsu, M.J. Buehler, DyFraNet: Forecasting and backcasting dynamic fracture mechanics in space and time using a 2D-to-3D deep neural network, APL Mach. Learn. 1 (2) (2023).

[21] E. Samaniego, C. Anitescu, S. Goswami, V.M. Nguyen-Thanh, H. Guo, K. Hamdia, X. Zhuang, T. Rabczuk, An energy approach to the solution of partial differential equations in computational mechanics via machine learning: Concepts, implementation and applications, Comput. Methods Appl. Mech. Eng. 15 (362) (2020) 112790.

[22] Zhang, L., You, H., Gao, T., Yu, M., Lee, C.H., Yu, Y., 2023. MetaNO: How to Transfer Your Knowledge on Learning Hidden Physics. arXiv preprint arXiv:2301.12095.

[23] M. Yin, E. Zhang, Y. Yu, G.E. Karniadakis, Interfacing finite elements with deep neural operators for fast multiscale modeling of mechanics problems, Comput. Methods Appl. Mech. Eng. 1 (402) (2022) 115027.

[24] E. Haghighat, M. Raissi, A. Moure, H. Gomez, R. Juanes, A physics-informed deep learning framework for inversion and surrogate modeling in solid mechanics, Comput. Methods Appl. Mech. Eng. 1 (379) (2021) 113741.

[25] He, J., Koric, S., Kushwaha, S., Park, J., Abueidda, D., Jasiuk, I., 2023. Novel DeepONet architecture to predict stresses in elastoplastic structures with variable complex geometries and loads. arXiv preprint arXiv:2306.03645.

[26] Z. Yang, C.-H. Yu, K. Guo, M.J. Buehler, End-to-end deep learning method to predict complete strain and stress tensors for complex hierarchical composite microstructures, J. Mech. Phys. Solids (2021), https://doi.org/10.1016/j.jmps.2021.104506.

[27] Z. Yang, C.-H. Yu, M.J. Buehler, Deep learning model to predict complex stress and strain fields in hierarchical composites, Sci. Adv. 7 (2021) eabd7416.

[28] Z. Yang, M.J. Buehler, Fill in the Blank: Transferrable Deep Learning Approaches to Recover Missing Physical Field Information, Adv. Mater. 35 (2023) 2301449, https://doi.org/10.1002/adma.202301449.

[29] Dall'Ara, E., Palanca, M., 2021. Data for paper "MicroFE models of porcine vertebrae with induced bone focal lesions: validation of predicted displacements with Digital Volume Correlation". The University of Sheffield. Dataset. ⟨https://doi.org/10.15131/shef.data.16732441⟩.

[30] M. Palanca, G. De Donno, E. Dall'Ara, A novel approach to evaluate the effects of artificial bone focal lesion on the three-dimensional strain distributions within the vertebral body, PLOS ONE 16 (6) (2021) e0251873, https://doi.org/10.1371/journal.pone.0251873.

[31] Rasband, W.S., ImageJ, U.S. National Institutes of Health, Bethesda, Maryland, USA, ⟨https://imagej.nih.gov/ij/⟩, 1997-2018.

[32] Stamati, et al., spam: Software for Practical Analysis of Materials, J. Open Source Softw. 5 (51) (2020) 2286, https://doi.org/10.21105/joss.02286.

[33] A. Sensini, O. Stamati, G. Marchiori, N. Sancisi, C. Gotti, G. Giavaresi, L. Cristofolini, M.L. Focarete, A. Zucchelli, G. Tozzi, Full-field strain distribution in hierarchical electrospun nanofibrous Poly-L(lactic) acid and Collagen based scaffolds for tendon and ligament tissue regeneration: a multiscale study, Heliyon (2024), https://doi.org/10.1016/j.heliyon.2024.e26796.

[34] M. Pena Fernandez, A. Barber, G. Blunn, G. Tozzi, Optimization of digital volume correlation computation in SR-microCT images of trabecular bone and bone-biomaterial systems (DOI: https://doi.org/), J. Microsc. (2018), https://doi.org/10.1111/jmi.12745.

[35] E. Dall'Ara, D. Barber, M. Viceconti, About the inevitable compromise between spatial resolution and accuracy of strain measurement for bone tissue: a 3D zero-strain study, J. Biomech. 47 (2014) 2956–2963, https://doi.org/10.1016/j.jbiomech.2014.07.019.

[36] G. Tozzi, E. Dall'Ara, M. Palanca, M. Curto, F. Innocente, L. Cristofolini, Strain uncertainties from two digital volume correlation approaches in prophylactically augmented vertebrae: local analysis on bone and cement-bone microstructures, JMBBM 67 (2017) 117–126.

[37] M. Palanca, L. Cristofolini, E. Dall'Ara, M. Curto, F. Innocente, V. Danesi, G. Tozzi, Digital volume correlation can be used to estimate local strains in natural and augmented vertebrae: An organ-level study, J. Biomech. 49 (2016) 3882–3890.

[38] Simonyan, K., Zisserman, A., 2014. Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556.

[39] Abadi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z., Citro, C., Corrado, G.S., Davis, A., Dean, J., Devin, M., Ghemawat, S., 2016. Tensorflow: Large-scale machine learning on heterogeneous distributed systems. arXiv preprint arXiv:1603.04467.

[40] I.S. Zhelavskaya, R. Vasile, Y.Y. Shprits, C. Stolle, J. Matzka, Systematic analysis of machine learning and feature selection techniques for prediction of the Kp index, Space Weather 17 (2019) 1461–1486, https://doi.org/10.1029/2019SW002271.

[41] K. Madi, G. Tozzi, Q.H. Zhang, J. Tong, A. Cossey, A. Au, et al., Computation of full-field displacements in a scaffold implant using digital volume correlation and finite element analysis, Med. Eng. Phys. 35 (2013) 1298–1312, https://doi.org/10.1016/j.medengphy.2013.02.001.

[42] G. Tozzi, Q.H. Zhang, J. Tong, Microdamage assessment of bone-cement interfaces under monotonic and cyclic compression, J. Biomech. 47 (2014) 3466–3474, https://doi.org/10.1016/j.jbiomech.2014.09.012.

[43] G. Tozzi, V. Danesi, M. Palanca, L. Cristofolini, Elastic full-field strain analysis and microdamage progression in the vertebral body from digital volume correlation (DOI:), Strain (2016), https://doi.org/10.1111/str.12202.

[44] H.H. Bayraktar, E.F. Morgan, G.L. Niebur, G.E. Morris, E.K. Wong, T.M. Keaveny, Comparison of the elastic and yield properties of human femoral trabecular and cortical bone tissue, J. Biomech. 37 (2004) 27–35.

[45] M. Rezasefat, H. Li, J.D. Hogan, Prediction of microstructural-dependent mechanical properties, progressive damage, and stress distribution from X-ray computed tomography scans using a deep learning workflow, Comput. Methods Appl. Mech. Eng. 424 (2024) 116878.