THESIS FOR THE DEGREE OF LICENTIATE OF ENGINEERING

Towards machine learning accelerated modelling of orthotropic composite materials

JOHAN FRIEMANN

Department of Mechanical Engineering
Chalmers University of Technology
Gothenburg, Sweden, 2026

Towards machine learning accelerated modelling of orthotropic composite materials

Johan Friemann

Technical Report No. IMS-2026-2
This thesis has been prepared using LaTeX.

Department of Mechanical Engineering
Chalmers University of Technology
SE-412 96 Gothenburg, Sweden
Phone: +46 (0)31 772 1000
`www.chalmers.se`

Printed by Chalmers Digitaltryck
Gothenburg, Sweden, April 2026

亜美奈と凱へ

Abstract

Laminated unidirectional composites and random short fibre composites have been homogenised with analytical methods for engineering design for decades. However there are no general analytical methods for woven (particularly 3D-reinforced) composites. There has been extensive work on the prediction of the elastic properties of 3D-reinforced textile composites through process modelling, but lately it has become popular to infer the as-manufactured geometry from X-ray computed tomography (XRCT). To derive the meso scale geometry from a XRCT volume the individual material phases must be segmented, which is a challenging problem. In this work the gap in the literature is being addressed by investigating machine learning (ML) based segmentation methods.

The first paper addresses the challenge of XRCT image segmentation, and more precisely the need for annotated data for training of ML based segmentation models. A fully automated pipeline for the generation of synthetic, automatically labelled, training data for ML based segmentation models of XRCT scans of 3D-textile reinforced composites is presented. A segmentation model trained on a dataset utilising the proposed pipeline shows 88% pixel-wise agreement when compared with a hand-segmented XRCT scan.

The second paper expands on the pipeline presented in the first paper with the aim to perform elastic homogenisation of a 3D-textile reinforced composite sample. The segmentation model is upgraded to a 3D U-Net architecture. The pipeline is also expanded to create voxel meshes from the segmentations including periodic boundary condition application. A novel material mapping routine is implemented that allows the mapping of the local yarn orientation and fibre volume fraction to the voxel elements, even when it is not fully possible to separate individual yarns. Combining the ML segmentation model with the material mapping routine, finite element models are generated which can be used to predict the statistical variation of elastic properties through computational homogenisation. An application of the pipeline results in accurately predicted homogenised elastic stiffnesses, with a deviation from experiments of less than 6.5%.

Keywords: Orthotropic materials, carbon fibre reinforced polymers, X-ray computed tomography, segmentation, machine learning, finite element modelling, 3D-textile reinforced composites.

List of Publications

This thesis is based on the following publications:

A

Johan Friemann, Lars P. Mikkelsen, Carolyn Oddy, Martin Fagerström, "Synthetic, automatically labelled training data for machine learning based X-ray CT image segmentation: Application to 3D-textile carbon fibre reinforced composites". *Composites Part B: Engineering* 305 (2025), 112656, https://doi.org/10.1016/j.compositesb.2025.112656.

B

Johan Friemann, Carolyn Oddy, Lars P. Mikkelsen, Martin Fagerström, "From X-ray CT to finite element models: A fully automated pipeline for mesoscale modelling of as-manufactured textile composites". *Composites Science and Technology* 278 (2026), 111561, https://doi.org/10.1016/j.compscitech.2026.111561.

Acknowledgments

The author acknowledges valuable support and encouragement from his PhD supervisors Prof. Martin Fagerström, Dr. Carolyn Oddy, and Assoc. Prof. Lars P. Mikkelsen. The author also acknowledges support and encouragement from his colleagues at the division of Computational Mechanics and Materials Engineering.

The author acknowledges funding from Horizon Europe through the MSCA Doctoral Network RELIANCE: Reliance: REaL-tIme characterisation of ANisotropic Carbon-based tEchnological fibres, films and composites, grant no. 101073040.

Acronyms

XRCT:	X-ray computed tomography
CFRP:	Carbon fibre reinforced polymers
FE:	Finite element
FBP:	Filtered back projection
FDK:	Feldkamp–Davis–Kress
RTM:	Resin transfer moulding
RVE:	Representative volume element
UC:	Unit cell
ML:	Machine learning
CNN:	Convolutional neural network

Contents

I Overview 1

Part I

Overview

CHAPTER 1

Background and theory

1.1 Introduction

Carbon fibre reinforced polymer (CFRP) composites with an orthotropic macro scale response are an important class of engineering materials. To begin with, a designer can tune the stiffnesses in potentially three orthogonal directions, enabling optimisation of lightweight structures. Secondly, the lack of normal-shear coupling when loaded in the principal material directions, reduces the risk of high shear stresses arising during loading. This is important as large shear stresses typically lead to the initiation of plasticity (or failure/damage). A few notable examples are random short fibre composites, specially orthotropic laminate composites, as well as most textile composites. See Figure 1.1 for schematic illustrations. The most classical example are balanced and symmetric laminated composites. They have widely been used in high performance engineering structures. Random short fibre composites are important in the context of recyclability. When continuous laminated materials are recycled the fibres typically break or degrade. Chopped fibres are therefore used to manufacture sheet moulding compound, or used during injection moulding to create structural components.

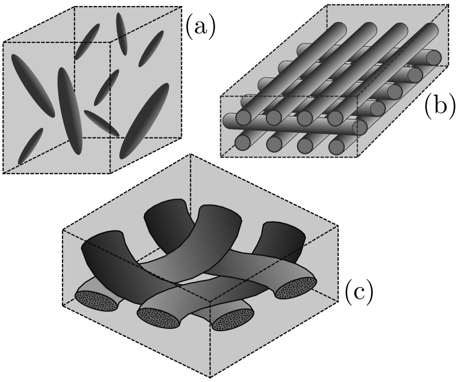


Figure 1.1: Schematics of (a) a random short fibre composite, (b) a symmetric cross-ply laminate, and (c) a 2D plain weave fabric are displayed. Both (a) and (b) are resolved at the micro scale while (c) resolves the meso scale.

Multi-layered textiles with a significant amount of yarns reinforcing the out-of-plane direction are called 3D-textile reinforced. When compared to laminated composites they have superior resistance for out-of-plane failure modes. Laminated composites are always facing the risk of delamination, where the each lamina is only connected to its neighbours by a pure matrix region. The added out-of-plane strength is particularly desirable in the aerospace industry where bird strikes, or turbine fan blade-offs can cause large out-of-plane impact loads. Commercially, a 3D-woven CFRP is already utilised in the LEAP engine fan blades and casing [1]. When compared to random short fibre composites, 3D-textile reinforced composites offer far higher fibre volume fractions and strength due to their continuous fibres. A notable type of 3D-textile reinforced material is the layer-to-layer angle interlock pattern. A layer-to-layer material consists of two yarn types (weft and warp[1]) like a standard 2D-textile as the one in Figure 1.1. The out-of-plane reinforcement is provided by yarns connecting several adjacent layers, which can be seen in the down-up-up-down weft pattern in the Figure 1.2.

During engineering design of load bearing composite structures, it is too

---

[1] Warp yarns are held under tension and extends vertically out of the loom in the weaving direction. Weft yarns are inserted over and under the warp yarns in a weave pattern, in the horizontal direction, creating the fabric.

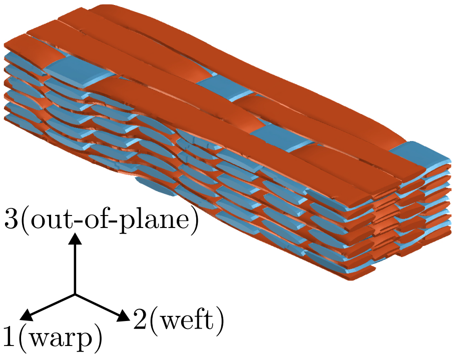


Figure 1.2: A schematic of a unit cell of a layer-to-layer angle interlock textile preform is displayed with the material coordinate system. Original graphics provided by courtesy of Oddy [2]. This material is studied in **Paper A** and **Paper B**.

costly to resolve the micro scale (individual fibres) or the meso scale (individual tows or yarns) when performing numerical simulations of the response of structural parts. A designer needs some type of micro (or meso) mechanical model to predict the elastic response of a part given the constituent properties and the micro (or meso) structure of the utilised material. Laminated composites are easily described in engineering design by classical laminate theory, and injection moulded short fibre composites are widely modelled with mean field theory (like the theory of Mori and Tanaka [3]) with orientation tensors as introduced by Advani et al. [4]. One drawback of 3D-textile reinforced composites is the difficulty in predicting the macro scale as-manufactured properties from the constituent properties and textile architecture. They typically need to be characterised through costly testing campaigns, or through finite element (FE) simulations. It may seem straightforward to predict the properties through computational homogenisation. However, the discrepancy between idealised computational geometries and the actual as-manufactured morphology makes prediction difficult.

The generation of representative volume elements (RVE) for computational homogenisation of 3D-textile reinforced composites is a widely studied topic. The most straight forward approach is pure geometric modelling like that done with tools like TexGen [5] or WiseTex [6]. As mentioned, the discrepancy between these geometries and the real as-manufactured textile shape can be significant. Therefore there has been extensive research in the so called digital element method where the as-manufactured geometry is predicted through a simulation of the yarn weaving method [7–12]. One issue with this approach is that the computational time per predicted unit cell (UC) can be very long.

A recent trend is to use X-ray computed tomography (XRCT) to scan as-manufactured material samples and use this information to derive computational geometry for RVE simulations [13–15]. It has been demonstrated that computational geometry derived from XRCT can significantly improve simulation accuracy when compared to geometry derived from other means [16]. One challenge in the use of XRCT is to consistently derive the computational geometry, namely image segmentation, local fibre volume fraction mapping, and fibre orientation mapping. This is difficult to automate, especially if the imaging contrast is poor. For certification purposes it is further motivated to be able to give statistical measures on structural as-manufactured parts. Therefore it would be ideal to be able to segment a large field of view scan encompassing several RVEs, like the scan shown in Figure 1.3. This in combination with the similarity of X-ray attenuation between carbon fibres and carbon based epoxy, makes poor contrast a tangible problem.

Machine learning (ML) based segmentation methods have become increasingly popular for segmentation of woven composite scans, and are seeing widespread use [17–19]. These methods show great promise and have the ability to outperform classical image segmentation techniques for poor contrast images. The main issue with ML based models however is that they shift the goal posts in automation. Most traditional ML based segmentation methods needs manually labelled data for training the model. This both creates the need for extensive manual work, and at the same time may lead to restrictions on the size of the training data set, potentially leading to overfitting. There is a clear gap in the literature in automated generation and labelling of training data for segmentation of XRCT images of 3D-textile reinforced composites. There has been some work in the use of synthetic automatically labelled data [20, 21]. There is however a further gap in the literature on basing the synthetic data on physical first principles.

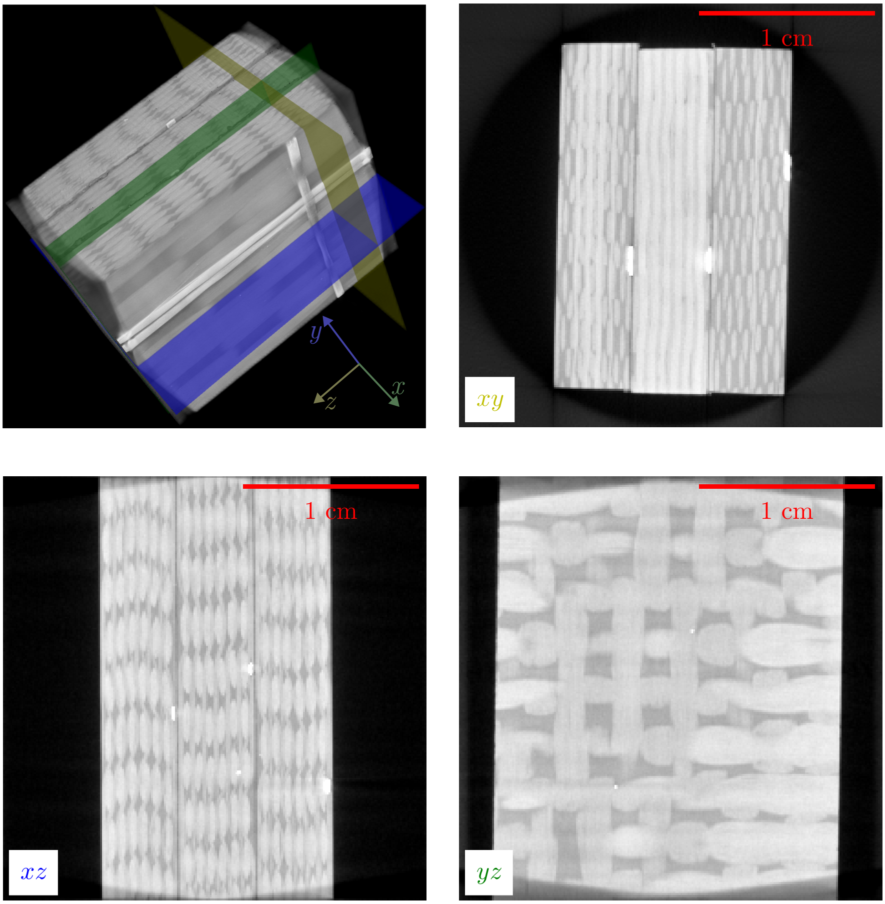


Figure 1.3: A 3D view of the XRCT scan of the 3D-woven sample studied in **Paper A** and **Paper B** is displayed together with three representative slices. The very bright areas in xy and xz are due to glass fibre tracer yarns, which are also visible on the surface in the 3D view.

1.2 Goals and scope of research

The aim of this research project is to use ML to accelerate the modelling of orthotropic composite materials. In the first stage of this project the prediction of the mechanical properties of 3D-textile reinforced composites is studied to encourage their further adoption in lightweight load bearing structures in the aerospace industry. In particular, ML based segmentation and subsequent elastic property homogenisation of 3D-textile reinforced textile composites is studied. The feasibility of utilising entirely synthetic automatically labelled training data is investigated in particular. To allow an algorithm trained on synthetic data to generalise to real data, the training data needs to be realistic and varied enough to bridge the so-called reality gap (or sim-to-real gap) [22]. In the context of this project, realistic simulations that can recreate the noise and artifacts one expects to see in an XRCT scan are performed. In the name of public interest, all novel aspects of the work are to be based on open source software and accessible to the wider research community without restrictions. The current goals and scope can be summarised as:

- SYNTHETIC DATA Create a pipeline that automatically generates labelled training data through realistic first principles simulations of XRCT scans.

- SIM-TO-REAL Demonstrate the ability of a segmentation algorithm trained on such data to segment real XRCT scans of large field of view scans of 3D-textile reinforced composites.

- HOMOGENISATION Utilise the verified pipeline to perform elastic homogenisation from many UCs from a single scan, and derive statistically significant engineering stiffness parameters.

- LIMITATION Limit analysis to small deformations and linear elasticity. As of now the study is limited to one material architecture.

1.3 Theory

Due to the interdisciplinary nature of this research project an extended theory section follows.

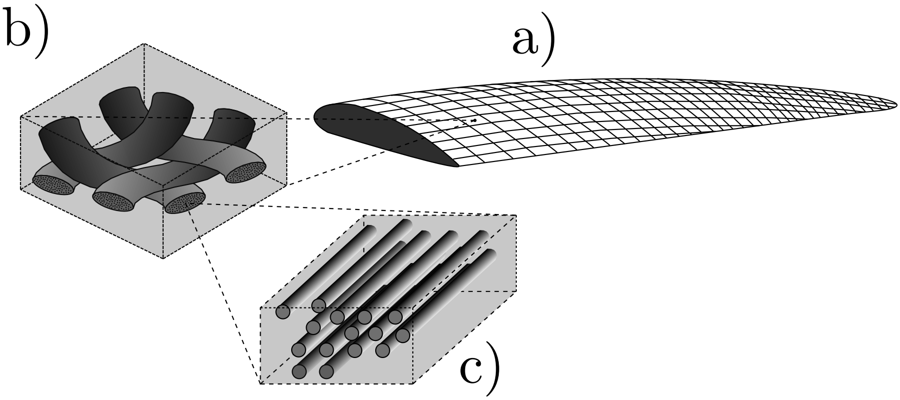


Figure 1.4: The figure illustrates the separation of scales in homogenisation. (a) The mesh of an airfoil constitutes the macro scale. (b) An integration point in the mesh encapsulates a unit cell of a textile, which represents the meso scale. (c) The micro scale resolves individual fibres, and determines the local material response in (b).

Computational homogenisation

It would be infeasible to resolve the individual carbon fibres in a CFRP structural part during engineering FE simulations. Homogenisation is an approach to create a surrogate continuum model for the macro scale response. As mentioned there are analytical models for the elastic response of continuous laminates. However for most materials homogenisation is performed computationally. The most important assumption is the separation of scale. That is, the length scale of the micro scale (or meso scale) is far smaller than the length scales considered in the macroscopic simulation. Refer to Figure 1.4 for a visualisation. Under such an assumption it is reasonable to decompose deformations into a macroscopic deformation, whose gradient does not vary across the micro scale, and a local fluctuation due to the local material inhomogeneity.

Let Ω represent the domain of an RVE. The average strain and stress over the RVE is defined as

$$\overline{\boldsymbol{\varepsilon}} = \frac{1}{|\Omega|} \int_{\Omega} \boldsymbol{\varepsilon} \, \mathrm{d}\Omega, \quad \overline{\boldsymbol{\sigma}} = \frac{1}{|\Omega|} \int_{\Omega} \boldsymbol{\sigma} \, \mathrm{d}\Omega. \tag{1.1}$$

The displacement field can be decomposed as

$$\boldsymbol{u}(\boldsymbol{x}) = \overline{\boldsymbol{\varepsilon}} \cdot \boldsymbol{x} + \tilde{\boldsymbol{u}}(\boldsymbol{x}), \tag{1.2}$$

where $\tilde{\boldsymbol{u}}(\boldsymbol{x})$ is the local fluctuation on the subscale. If the scales are truly separate the local variations should not directly affect the macroscopic response, and it should be possible to formulate

$$\overline{\boldsymbol{\sigma}} = \overline{\boldsymbol{C}} : \overline{\boldsymbol{\varepsilon}} \tag{1.3}$$

for some effective linear elastic constitutive relationship $\overline{\boldsymbol{C}}$. The homogenisation problem entails finding this effective constitutive relationship. It seems like it possible to simply perform an FE simulation over the RVE and use Equations (1.1) and (1.3) to determine the macroscopic response. However, the average is not guaranteed to yield unique results unless the boundary conditions of the simulation are chosen carefully.

The Hill-Mandel condition states that the work performed by the effective response should equal the average work over the subscale. The condition ensures energy equivalent across scales, that is,

$$\overline{\boldsymbol{\sigma}} : \overline{\boldsymbol{\varepsilon}} = \frac{1}{|\Omega|} \int_\Omega \boldsymbol{\sigma} : \boldsymbol{\varepsilon} \, \mathrm{d}\Omega. \tag{1.4}$$

Boundary conditions that satisfy the Hill-Mandel conditions are guaranteed to result in a symmetric and positive definite $\overline{\boldsymbol{C}}$.

It turns out that periodic displacement boundary conditions satisfies the Hill-Mandel condition. For small strains

$$\begin{aligned}&\int_\Omega \boldsymbol{\sigma} : \boldsymbol{\varepsilon} \, \mathrm{d}\Omega = \int_\Omega \boldsymbol{\sigma} : \boldsymbol{\nabla} \boldsymbol{u} \, \mathrm{d}\Omega = \int_\Omega \boldsymbol{\nabla}(\boldsymbol{\sigma} \cdot \boldsymbol{u}) \, \mathrm{d}\Omega - \int_\Omega \boldsymbol{\nabla} \cdot \boldsymbol{\sigma} \cdot \boldsymbol{u} \, \mathrm{d}\Omega = \\ &(\text{Div. Th. and } \boldsymbol{\nabla} \cdot \boldsymbol{\sigma} = \boldsymbol{0}) = \int_\Gamma \boldsymbol{t} \cdot \boldsymbol{u} \, \mathrm{d}\Gamma,\end{aligned} \tag{1.5}$$

where the traction is defined as $\boldsymbol{t} = \boldsymbol{n} \cdot \boldsymbol{\sigma}$. By using the decomposition in Equation (1.2) it follows that

$$\int_\Gamma \boldsymbol{t} \cdot \boldsymbol{u} \, \mathrm{d}\Gamma = \int_\Gamma \boldsymbol{t} \cdot (\overline{\boldsymbol{\varepsilon}} \cdot \boldsymbol{x} + \tilde{\boldsymbol{u}}) \, \mathrm{d}\Gamma = \int_\Gamma \boldsymbol{t} \otimes \boldsymbol{x} \, \mathrm{d}\Gamma : \overline{\boldsymbol{\varepsilon}} + \int_\Gamma \boldsymbol{t} \cdot \tilde{\boldsymbol{u}} \, \mathrm{d}\Gamma. \tag{1.6}$$

If individual RVEs are to be in equilibrium the traction $\boldsymbol{t}$ must be anti-periodic. As stated above, there are periodic displacement boundary conditions. As a consequence the second term in the far RHS in Equation (1.6) thus disappears. Now use the divergence theorem again backwards to get

$$\begin{gathered}\int_{\Gamma} \boldsymbol{t} \otimes \boldsymbol{x} \, \mathrm{d}\Gamma : \bar{\boldsymbol{\varepsilon}} = \int_{\Omega} \boldsymbol{\nabla} \cdot (\boldsymbol{\sigma} \otimes \boldsymbol{x}) \, \mathrm{d}\Omega : \bar{\boldsymbol{\varepsilon}} = \\ \int_{\Omega} (\boldsymbol{\nabla} \cdot \boldsymbol{\sigma}) \otimes \boldsymbol{x} \, \mathrm{d}\Omega + \int_{\Omega} \boldsymbol{\sigma} \, \mathrm{d}\Omega : \bar{\boldsymbol{\varepsilon}} = |\Omega| \bar{\boldsymbol{\sigma}} : \bar{\boldsymbol{\varepsilon}}.\end{gathered} \tag{1.7}$$

This is the Hill-Mandel condition. This demonstrates that periodic displacement boundary conditions are suitable for FE RVE homogenisation. It is also possible to use appropriately selected Dirichlet or Neumann conditions. It can be shown that they provide upper (Dirichlet conditions, Voigt assumption) and lower (Neumann conditions, Reuss assumption) bounds on the effective stiffness. The periodic boundary condition stiffness must therefore be in-between and typically converges quicker for increasing RVE sizes[2]. It is the most suitable and accurate boundary condition if the material is fully periodic and the RVE encompasses one (or several) UCs.

Orthotropic materials are the most general form of anisotropic materials that do not display deformation normal-shear coupling in the frame of reference aligned with the main material axes. Any orthotropic linearly elastic material can be described by 9 independent parameters. Consider an orthotropic elastic material, where the material directions $1, 2, 3$ are ordered from stiffest to most compliant. For such a material there are three Young's moduli (E_{11}, E_{22}, and E_{33}), three major Poisson's ratios (ν_{12}, ν_{23}, and ν_{13}), and three shear moduli (G_{12}, G_{23}, and G_{13}). The Voigt form of the stress strain relationship of a linearly elastic orthotropic material is given by

$$[\boldsymbol{\varepsilon}]=[\boldsymbol{S}][\boldsymbol{\sigma}]=\begin{bmatrix}\varepsilon_{11}\\ \varepsilon_{22}\\ \varepsilon_{33}\\ \gamma_{12}\\ \gamma_{23}\\ \gamma_{13}\end{bmatrix} = \begin{bmatrix}\frac{1}{E_{11}} & -\frac{\nu_{12}}{E_{11}} & -\frac{\nu_{13}}{E_{11}} & 0 & 0 & 0\\ -\frac{\nu_{12}}{E_{11}} & \frac{1}{E_{22}} & -\frac{\nu_{23}}{E_{22}} & 0 & 0 & 0\\ -\frac{\nu_{13}}{E_{11}} & -\frac{\nu_{23}}{E_{22}} & \frac{1}{E_{33}} & 0 & 0 & 0\\ 0 & 0 & 0 & \frac{1}{G_{12}} & 0 & 0\\ 0 & 0 & 0 & 0 & \frac{1}{G_{23}} & 0\\ 0 & 0 & 0 & 0 & 0 & \frac{1}{G_{13}}\end{bmatrix}\begin{bmatrix}\sigma_{11}\\ \sigma_{22}\\ \sigma_{33}\\ \sigma_{12}\\ \sigma_{23}\\ \sigma_{13}\end{bmatrix}. \tag{1.8}$$

---

[2] If separation of scale holds the effective response should not change noticeably for an RVE containing one or several unit cells.

If 6 stress controlled[3] FE simulations are carried out where only one of the components of $\boldsymbol{\sigma}$ is non-zero per simulation, all the independent material parameters can be obtained. It is easy to see that for each load case, the output $\boldsymbol{\varepsilon}$ scaled by the non-zero input control stress yields a column of $\boldsymbol{S}$.

In the case of a meso scale RVE, like a woven material, the yarns themselves need to be homogenised before an RVE simulation can be carried out. This would correspond to go from (c) to (b) in Figure 1.4. This could be done with another FE model, but another common approach is to use an analytical micromechanical model. This approach was adopted in **Paper B**. For such a two-step homogenisation, each yarn element (or integration point) in a simulation of (b) would require a local fibre volume fraction, and a local fibre orientation.

In order for the homogenisation to yield representative macroscopic properties, the RVE geometry and the constituent properties need to be representative and accurate. As mentioned, using purely geometric modelling yields a large discrepancy when compared to the actual as-manufactured material morphology. It is therefore of interest to use computed tomography to infer the as-manufactured geometry.

Computed tomography

By exploiting the differences in energy attenuation of X-rays between the different material phases, it is possible to reconstruct the internal micro (or meso) structure of an as-manufactured composite sample. X-rays are emitted from an X-ray source, pass through the sample under study, and the remaining X-ray energy is measured at a detector. The detector consists of an array of imaging pixels (like a digital camera). The resulting images are referred to as X-ray projections, and describes the sum of all attenuation through the entire sample. The computed tomography problem entails taking several projections from different angles and reconstructing the internal geometry.

The computed tomography problem is stated here for a parallel X-ray beam [4] for ease of demonstration. Consider a body transmissible to X-rays placed at the centre of a Cartesian coordinate system. An X-ray source emitting

---

[3] A similar argument can be made with strain-controlled simulations, where ε is an input and σ is an output.

[4] For parallel beams it holds that any X-ray that hits a detector pixel at height z will only ever pass through material in the sample at that same height. It then follows that reconstruction can be performed independently for each z.

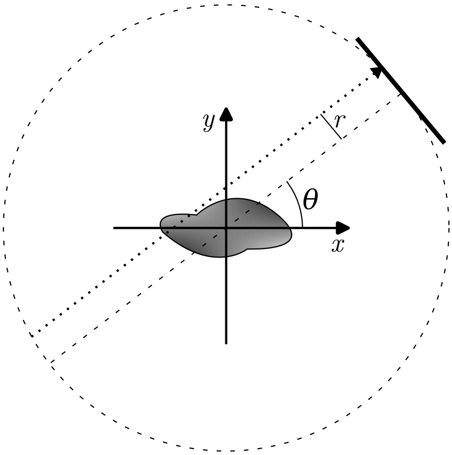


Figure 1.5: The projection corresponding to a rotation θ, of a ray (dotted arrow) at distance r away from the detector centre (dashed line).

parallel rays, and a detector is placed at a fixed distance from the origin of a Cartesian coordinate system. They are allowed to rotate together around the centre of the coordinate system, as shown in Figure 1.5. Let $D_{r,\theta,z}$ be a ray at height z perpendicular to the detector at distance r away from the centre of the detector. Also let θ be the angle that the detector source pair is rotated around the origin[5]. For monochromatic rays, the intensity of X-rays passing through the body along $D_{r,\theta,z}$ is given by the Beer-Lambert law:

$$I(r,\theta,z) = I_0 \exp\left(-\int_{D_{r,\theta,z}} \mu(x,y,z)\ \mathrm{d}s\right), \tag{1.9}$$

where I_0 is the intensity of the X-rays emitted by the source, and μ is an attenuation coefficient. The attenuation coefficient depends on the atomic number of the material, with higher number elements displaying higher attenuation. From Equation (1.9) it immediately follows that

$$\int_{D_{r,\theta,z}} \mu(x,y,z)\ \mathrm{d}s = -\ln\left(\frac{I}{I_0}\right) := p(r,\theta,z), \tag{1.10}$$

---

[5] The setup outlined here is equivalent with holding the detector and source fixed and rotating the body.

where the function $p(r,\theta)$ represents a *flat field corrected* tomographic projection of the body rotated by an angle θ with detector coordinate r. The tomographic reconstruction problem consists of finding the attenuation coefficients μ given the set of projections $p(r,\theta)$. By using the Dirac delta functional Equation (1.10) is rewritten as

$$p(r,\theta,z) = \iint_{\mathbb{R}^2} \mu(x,y,z)\delta(x\cos\theta + y\sin\theta - r)\ \mathrm{d}x\ \mathrm{d}y. \tag{1.11}$$

It is possible to show from the central slice theorem that

$$\mu(x,y,z) = \int_0^\pi p(r,\theta,z) * b(r)\ \mathrm{d}\theta, \tag{1.12}$$

where $(*)$ is the convolution operator and b is a ramp filter with Fourier transform $\mathcal{F}\{b(r)\}(\omega) = |\omega|$. This is the filtered back projection (FPB) algorithm. In practice, the integral in Equation (1.12) becomes a sum, as it is not possible to take an infinite amount of tomographic projections. The discretised version will introduce noise and artifacts that may make image contrast poorer. Furthermore, the limited resolution of the detector discretises $p(r,\theta)$, further degrading reconstruction quality. The resulting reconstructed volume will be a discretisation of μ where the size of voxels will depend on the size of the detector pixels.

In contrast to the idealised parallel beam case, lab based XRCT equipment typically utilises a cone beam geometry, like the one shown in Figure 1.6. One benefit of the cone beam geometry is the optical magnification provided by the diverging X-rays. Yet for such a set-up, the above derived reconstruction algorithm yields erroneous results. There is however a modified algorithm called the the Feldkamp-Davis-Kress (FDK) technique [23], which is a generalisation of FBP. It is worth mentioning that unlike FBP, that is exact in the limiting case of infinite projections, FDK is an approximate method. The approximations in standard implementations of FDK typically result in a longitudinal fall-off effect where densities are underestimated at the extreme ends of the vertical axis [24]. In this work, the FDK algorithm is utilised for reconstructing XRCT scans exclusively.

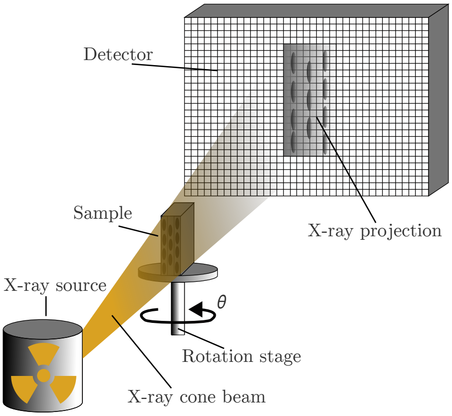


Figure 1.6: The figure shows a schematic representation of a cone beam lab XRCT set-up. Note the geometric magnification from the cone beam.

The segmentation problem

Given the attenuation coefficients as a function of the spatial coordinates $\mu(x, y)$, it is in theory possible to gain an understanding of a studied materials internal structure and separate the different material phases. In reality the limited number of projections, the limited resolution of the detector, electronic noise from the detector, and noise due to the quantum nature of light, the phases may be difficult to distinguish from each other. This is made worse by the fact that the constituents of CFRPs have very similar attenuation coefficients due to the similar chemical compositions of the material phases. Specifically, carbon fibres surrounded by an epoxy matrix which consists mostly of carbon itself.

The problem of assigning each pixel in an image (or voxel in a volume) to its corresponding material class is called segmentation. See Figure 1.7 for a schematic demonstration of an image segmentation. The segmentation is important for creating computational geometry for RVE simulations. In the material studied in **Paper A** and **Paper B** there are two different yarn types with a different number of filaments. Furthermore, individual yarns need to be separated in order to infer the local orientation of the filaments. The number of filaments will influence the local fibre volume fraction, and will thus result in a difference in contrast.

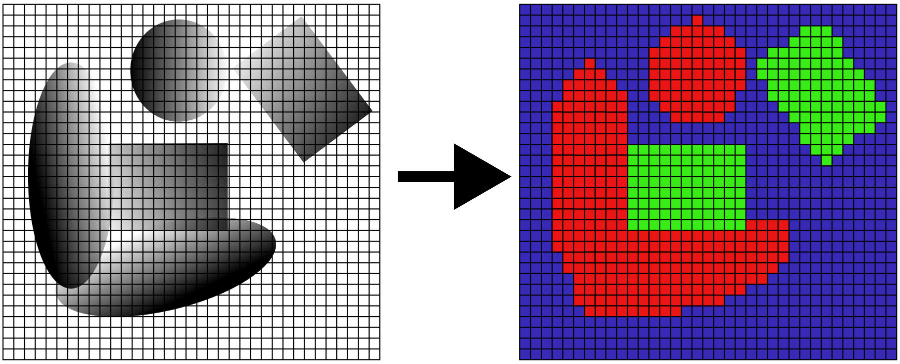


Figure 1.7: The figure shows a schematic example of an image segmentation. The greyscale intensities at each pixel on the left are assigned to one of three material classes on the right: red for round shapes, green for square shapes, and blue for background.

There are several classical methods that use the greyscale intensity information of an image (or volume). Either using global histograms like Otsu's method [25], or local intensity variations like watershed segmentation [26]. Owing to the particularly poor contrast of reconstructions of CFRP scans, classical methods like Otsu and watershed become infeasible [27, 28]. A visualisation of the shortcomings of Otsu thresholding for a textile reinforced CFRP are demonstrated in Figure 1.8. Note how the method struggles to distinguish between yarn types, and how it erroneously labels noise and artifacts as belonging to the yarn phase. Similarly to Otsu thresholding, watershed segmentation is also sensitive to noise. The results can be improved by using structural information in the image (cf. the streaks in Figure 1.8) through for example structure tensor analysis, as was done by Naouar et al. [29]. Structure tensor analysis can however also be difficult when the resolution is not high enough to resolve important structural features. This is the case for particularly large field of view scans, where the resolution (voxel size) becomes significantly larger than the fibre diameter. Cf. the scan in Figure 1.3, where the voxel size is approximately nine times larger than the fibre diameter.

It is possible for a human to segment the slice in Figure 1.8 without too much of effort. Humans can utilise contextual information, prior knowledge, and higher order features that a simple intensity based segmentation algorithm

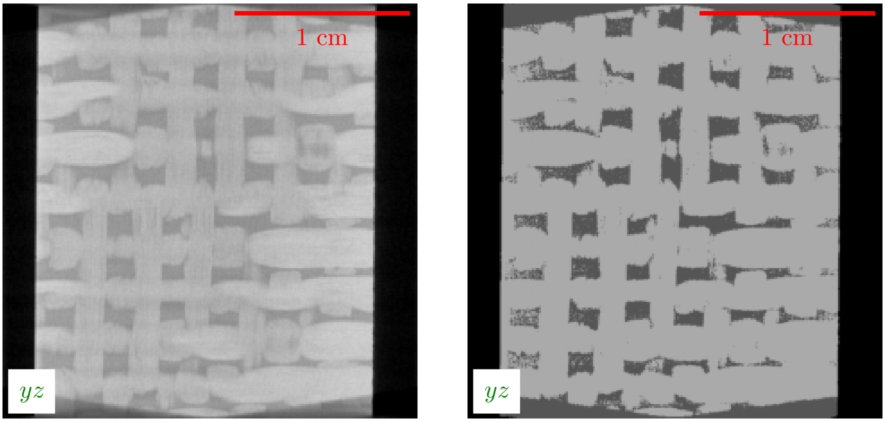


Figure 1.8: A slice of an XRCT scan studied in this work (left) shown next to a 4-class Otsu threshold segmentation (right). Note how Otsu is not able to differentiate between weft and warp yarns. Furthermore, noise gets categorised as yarns, and the cone beam artifact gets segmented as part of the matrix phase.

does not have access to. It is therefore motivated to investigate segmentation algorithms that are context aware, like ML based segmentation.

Image segmentation with artificial neural networks

Convolutional Neural Networks (CNN) can learn complex relationships in data and are promising for segmentation of XRCT images or volumes. When compared to fully connected artificial neural networks, CNNs can have more complex architecture without adding too much computational complexity due to convolution kernels sharing weights.

The CNN is here presented for 2D images for the sake of simplicity. Extension to 3D follows trivially through the addition of an index. Let I represent an image. Then, I_{cij} represents the value of channel c (for example RGB values of an image) of the pixel with coordinates i, j. Furthermore let $\hat{I}_{cij}$ represent the "activated" image passed through an activation function ϕ, like a Rectified Linear Unit (ReLU). Let K_{dcpq} be the weights of a c input channel d output channel convolution kernel of size $p \times q$, and let B_d be a d channel bias. Consider a multi layer convolutional neural network. Given an activated image of a layer $n-1$ the activated image output from layer n is given by (utilising Einstein summation convention)

$$I_{dij}^{(n)} = K_{dcpq}^{(n)} \hat{I}_{c(i+p)(j+q)}^{(n-1)} + B_d^{(n)}, \; \hat{I}_{dij}^{(n)} = \phi(I_{dij}^{(n)}). \tag{1.13}$$

The goal is for a given network architecture (no. layers, kernel size etc.) to choose the kernels such that when fed an image the output of the final layer becomes a segmented image. This is achieved through training, where the network is shown an image and the difference between the output and a segmented ground truth is used to compute an error. The error is minimised through taking a step in the gradient with respect to the kernels and biases.

For a network with N layers, the error is computed through a loss function $\mathcal{L}(\hat{I}^{(N)}, \hat{I}_{\text{truth}})$, where one example is the cross entropy loss. The process of computing the gradient of the loss with the respect to the kernel weights and biases is referred to as backpropagation and is here derived for the kernel weights. The derivative with respect to the biases are derived analogously.

Through the chain rule the gradient of the loss with respect to an arbitrary layer n's kernel weights becomes

$$\frac{\partial \mathcal{L}}{\partial K_{dcpq}^{(n)}} = \frac{\partial \mathcal{L}}{\partial \hat{I}_{d'i'j'}^{(n)}} \frac{\partial \hat{I}_{d'i'j'}^{(n)}}{\partial I_{d''i''j''}^{(n)}} \frac{\partial I_{d''i''j''}^{(n)}}{\partial K_{dcpq}^{(n)}} \tag{1.14}$$

It is evident from Equation (1.13) that the derivative of an output of the image of layer n with respect to the current layer's kernel becomes the previous layer's activation:

$$\frac{\partial I_{d'ij}^{(n)}}{\partial K_{dcpq}^{(n)}} = \hat{I}_{c(i+p)(j+q)}^{(n-1)} \delta_{d'd}, \tag{1.15}$$

where δ is the Kronecker delta. The second factor is the derivative of the activation functions. The first factor in Equation (1.14) can be computed for the output layer $n = N$. If a relationship between the derivatives of layer n and $n-1$ can be found, all the kernel derivatives may be determined through recursion, which constitutes the backpropagation algorithm. It follows from Equation (1.13) that

$$\frac{\partial I_{d'i'j'}^{(n)}}{\partial \hat{I}_{dij}^{n-1}} = K_{d'dpq}^{(n)} \delta_{(i'+p)i} \delta_{(j'+q)j} \tag{1.16}$$

Now assume the gradient of $\mathcal{L}$ with respect to I of the nth layer is known, then

$$\begin{aligned}\frac{\partial \mathcal{L}}{\partial \hat{I}^{(n-1)}_{dij}} &= \frac{\partial \mathcal{L}}{\partial I^{(n)}_{d'i'j'}} \frac{\partial I^{(n)}_{d'i'j'}}{\partial \hat{I}^{(n-1)}_{dij}} \\ \text{(using Eq. (1.16))} &= \frac{\partial \mathcal{L}}{\partial I^{(n)}_{d'i'j'}} K^{(n)}_{d'dpq} \delta_{i(i'+p)} \delta_{j(j'+q)} \\ &= \frac{\partial \mathcal{L}}{\partial I^{(n)}_{d'(i-p)(j-q)}} K^{(n)}_{d'dpq}.\end{aligned} \tag{1.17}$$

This is a convolution between the nth layers gradient with respect to the nth layers output and the nth layer's kernel rotated by 180 degrees. Using these results, the gradient for an arbitrary layer can be computed by repeated convolutions.

In practice CNNs consist of a few more building blocks. One important block is the max pool operation which functions like a convolution kernel that extracts the maximum value inside the sliding window. Max pooling reduces the spatial dimensions (potentially reducing required memory) while retaining the most important information in an image. It is differentiable in the context of backpropagation and its gradient will simply be an indicator function of the maximum value. Another common block is the up-convolution, where a transverse convolution operation with padding is utilised, making the output image larger than the input. The ability to up-sample the data is important for a segmentation algorithm, where the output image needs to have the same dimensions as the input.

One problem facing large CNNs is the vanishing gradient problem, where deep backpropagation makes the gradients smaller than machine precision. Therefore, complex segmentation networks can become difficult to train. The architecture needs to be modified to remedy this.

A common architecture for image segmentation is an encode-decoder type network with skip connections, like U-Net [30]. In a skip connection, the output of one layer in the encoder is concatenated to the input of some layer in the decoder. This amounts to adding a term depending on $\hat{I}$ from layer $(n-k)$ to Equation (1.13) [6]. It is clear that the backpropagation algorithm derived above generalises to this modification. The skip connections eases

---

[6] Formally new channels are added to $\hat{I}$, but it can be expressed as a new term where the channels concerning the other term are constant zero.

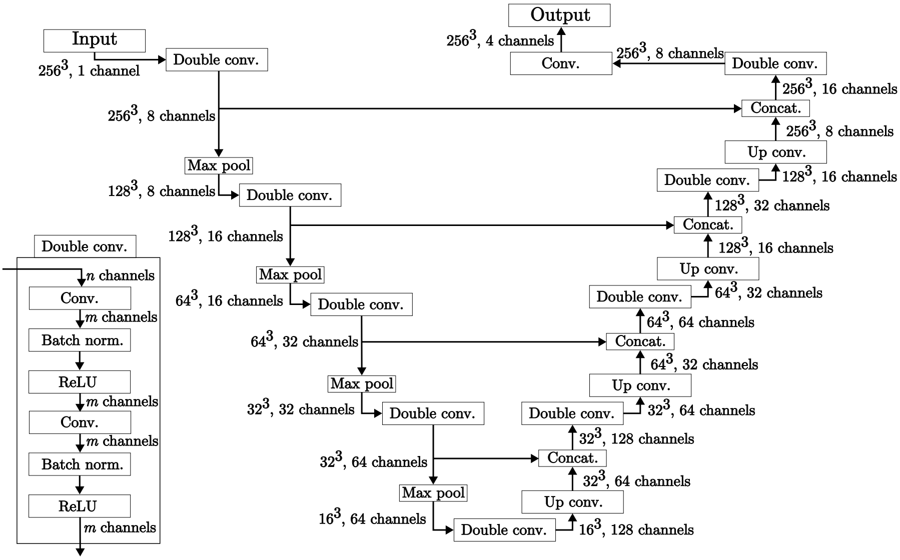


Figure 1.9: The 3D convolution U-Net architecture used in **Paper B** is displayed. The input is a one channel (greyscale) XRCT image and the output with four channels represents the segmentation class probabilities. The structure of the *Double convolution* block is also shown.

gradient flow through the network making the vanishing gradient problem less severe. See Figure 1.9 for a schematic of a U-Net architecture.

Optimisation of $\mathcal{L}$ is typically done with a variant of stochastic gradient descent. For each weight update the average error over a few data points is computed. This type of training usually requires a tremendous amount of labelled training data to not be overfit (akin to fitting an nth degree polynomial to $n + 1$ points). Procuring and labelling such quantities of training data can be incredibly time consuming and expensive, which is one of the main criticism of neural networks. If it is possible to utilise a completely synthetic data set (that is automatically labelled) this shortcoming can be circumvented.

CHAPTER 2

Summary of included papers

This chapter provides a summary of the included papers.

2.1 Paper A

Johan Friemann, Lars P. Mikkelsen, Carolyn Oddy, Martin Fagerström
Synthetic, automatically labelled training data for machine learning based X-ray CT image segmentation: Application to 3D-textile carbon fibre reinforced composites
Published in Composites Part B: Engineering,
Vol. 305, 112656, Oct. 2025.
©2025 The Authors. Published by Elsevier Ltd.
https://doi.org/10.1016/j.compositesb.2025.112656 .

In this paper a fully automated pipeline for generating synthetic labelled training data for ML based segmentation of 3D-textile reinforced composites is presented. The aim of the work is to provide a method to generate a tailor-made training data set for a material architecture and XRCT scanner set-up.

In contrast to earlier works in utilising synthetic training data [20, 21], the current work includes physically motivated noise and artifacts by performing a simulation of the XRCT scan. The pipeline is efficient enough to allow for generating hundreds of synthetic XRCT volumes with corresponding segmentations in a feasible time frame. The idea is to adapt the simulation to the sample to be segmented and the equipment used to scan the sample, and perform domain randomisation (randomly varying simulation parameters for different virtual samples to capture the 'real' parameters in the statistics) on a large data set, in order to aid in transferring a trained segmentation algorithm to an actual sample.

The first step of the pipeline entails generating geometry. 3D surface meshes suitable for ray-tracing simulations are generated with the textile modelling software TexGen. In order to bridge the reality gap through domain randomisation, a wide variety of weave architectures can be randomised. Furthermore, random perturbations and deformations are included in the yarns to make the geometry less idealised.

Secondly, the generated surface meshes are scanned virtually with a ray tracing simulation, and the resulting X-ray projections are reconstructed with the same algorithm used to reconstruct the real sample under study. The simulation accounts for the specific properties of the XRCT equipment used and the chemical composition of the sample constituents.

The third and final step entails performing a solid voxelisation of the virtually scanned surface meshes. A novel contribution of the present work was an adaptation of a voxelisation algorithm to perform the ground truth labelling in the same frame of reference as the reconstructed XRCT scan, yielding a one-to-one correspondence between the XRCT volume and the ground truth segmented voxels.

The viability of the presented pipeline is demonstrated by segmenting an actual large field of view XRCT scan of a layer-to-layer angle interlock composite sample. The segmentation is compared with a hand segmented ground truth, and an 88% pixel-wise agreement is obtained.

2.2 Paper B

Johan Friemann, Carolyn Oddy, Lars P. Mikkelsen, Martin Fagerström
From X-ray CT to finite element models: A fully automated pipeline for mesoscale modelling of as-manufactured textile composites
Published in Composites Science and Technology,
Vol. 278, 111561, May. 2026.
©2026 The Authors. Published by Elsevier Ltd.
https://doi.org/10.1016/j.compscitech.2026.111561 .

In this paper, the pipeline presented in **Paper A** is expanded and utilised to perform elastic property homogenisation of a layer-to-layer angle interlock CFRP sample. A large field of view scan is automatically segmented and a novel fibre volume fraction and local fibre orientation mapping procedure is implemented. The segmentation is performed with a new 3D U-Net architecture, where the synthetic fully labelled volumetric dataset enables utilising 3D-convolutions that provide out-of-plane context to the segmentation algorithm. The new segmentation architecture greatly improves yarn continuity in the segmentations when compared to the algorithm used in **Paper A**.

To generate a finite element discretisation of the segmented geometry, a voxel meshing approach is used, where each voxel in a UC corresponds to an element in the RVE simulation. It is possible to create 6 non-overlapping UCs from one scan, and the novel material mapping procedure is able to satisfactorily map local fibre orientation even when it is impossible to fully separate individual yarns in highly compacted regions of a UC.

The pipeline handles everything from segmentation, to meshing, to boundary condition assignment. The only manual labour required is the selection of the individual UCs from the large field of view XRCT segmentation. The proposed method allows for accurate prediction of the elastic homogenised stiffnesses of a 3D reinforced CFRP, with a deviation of less than 6.5% from experimentally measured stiffnesses.

CHAPTER 3

Future work

In order to expand the analysis to large deformations, plasticity, and damage, the number of elements in the RVE models need to be reduced. Therefore, conformal meshing is required. In addition to typically consisting of a much smaller number of elements, conformal meshes also do not display the sharp corners found in voxel meshes that can introduce unrealistic stress concentrations. Such stress concentrations can lead to premature plasticity and damage, yielding inaccurate strength predictions. In addition to expanding the study of 3D-textile reinforced composites, there are other gaps in the literature on orthotropic CFRPs where the combination of XRCT and ML based models trained synthetic data can be highly useful. One example is compressive failure via fibre kinking of laminated composites.

A large part of the work in **Paper B** was dedicated to finding a robust method for determining the local fibre orientation inside the reinforcing yarns. If the segmentation used for meshing had a clear separation between individual yarns, this task would have been trivial. As of now there is no straightforward way to guarantee well separated yarns in the segmentations studied in this work. An additional challenge is that the geometry used for the synthetic XRCT data was generated with a geometric modeller. Even though there are perturbations added, the geometry is still far too idealised. In particular there are no highly compacted regions where yarns are smoothly touching. Furthermore, TexGen has a tendency to produce self-intersecting geometry for highly compacted weaves. In addition, the yarns are further modelled as completely homogenous. It would be interesting to train a model to directly output the local fibre volume fraction and fibre orientation instead of inferring it from the segmentation. This would require a rework of the geometry stage of the pipeline.

The results so far indicate that the largest shortcoming of the current pipeline is the over idealised geometry for the XRCT simulation. It is likely that the simulation to real transferability could increase further if more realistic, compacted, yarns are generated. Furthermore, the introduction of inhomogeneous yarns could aid a segmentation algorithm to distinguish diffuse borders between yarns seen in actual XRCT scans. Therefore, for future improvements of the pipeline introduced in **Paper A** and refined in **Paper B**, it is of interest to replace the geometry generation method. An improved geometry generation method should include:

- LESS IDEALISED GEOMETRY A geometry generation method should be able to create less idealised, closer to as-manufactured geometry.
- ORIENTATION INFORMATION The geometry generation should also include information about local fibre orientation and local fibre volume fraction.
- CONFORMAL MESHING The method should be able to provide conformal meshes. It should further allow segmentation algorithms to separate yarns which are suitable for automated conformal meshing.
- INCREASED GENERALITY The geometry generation should not only be applicable to 3D-textile reinforced composite materials, but also other material architectures.

Based on these requirements it is tempting to return to yarn weaving (alternatively fibre pultrusion) process modelling to generate closer to reality as-manufactured geometry. As mentioned in the introduction, existing methods are computationally expensive and would thus not be suitable to generate a large synthetic training set. If it is possible to significantly speed up the process modelling to achieve real time simulation performance this approach would be very promising. Therefore, it is of interest to take inspiration from simulation techniques utilised in the computer graphics research community, where graphics processing units are used to perform massively parallel simulations. If the output of such simulations are used for synthetic training data, any loss of accuracy due to highly optimised approximate physics simulation, would be less critical than for direct prediction of mechanical properties. Gao et al. [31] used position based dynamics to post-process idealised geometry to increase realism. Inspired by this development, and the state of the art modelling approaches in computer graphics like vertex block descent [32, 33], a natural next step of this project is to implement a real time simulation process modeller for the geometry stage of the pipeline.

References

[1] M. N. Saleh and C. Soutis, “Recent advancements in mechanical characterisation of 3D woven composites,” *Mechanics of Advanced Materials and Modern Processes*, vol. 3, no. 1, Jul. 2017, ISSN: 2198-7874.

[2] C. Oddy et al., “Cyclic behaviour of 3D-woven composites in tension: Experimental testing and macroscale modelling,” *Composites Part A: Applied Science and Manufacturing*, vol. 187, p. 108 354, Dec. 2024, ISSN: 1359-835X.

[3] T. Mori and K. Tanaka, “Average stress in matrix and average elastic energy of materials with misfitting inclusions,” *Acta Metallurgica*, vol. 21, no. 5, pp. 571–574, May 1973, ISSN: 0001-6160.

[4] S. G. Advani and C. L. Tucker, “The use of tensors to describe and predict fiber orientation in short fiber composites,” *Journal of Rheology*, vol. 31, no. 8, pp. 751–784, Nov. 1987, ISSN: 1520-8516.

[5] L. P. Brown and A. C. Long, “Modeling the geometry of textile reinforcements for composites: TexGen,” in *Composite Reinforcements for Optimum Performance*, ser. Woodhead Publishing Series in Composites Science and Engineering, P. Boisse, Ed., Second Edition, Woodhead Publishing, 2020, pp. 237–265, ISBN: 9780128190050.

[6] S. V. Lomov, *Modeling the geometry of textile composite reinforcements: WiseTex*, 2nd ed. Elsevier Ltd., 2020, pp. 199–236, ISBN: 9780128190050.

[7] Y. Wang and X. Sun, “Digital-element simulation of textile processes,” *Composites Science and Technology*, vol. 61, no. 2, pp. 311–319, 2001, ISSN: 02663538.

[8] G. Zhou, X. Sun, and Y. Wang, “Multi-chain digital element analysis in textile mechanics,” *Composites Science and Technology*, vol. 64, no. 2, pp. 239–244, 2004, ISSN: 02663538.

[9] B. S. El Said, S. Green, and S. R. Hallett, “Kinematic modelling of 3D woven fabric deformation for structural scale features,” *Composites Part A: Applied Science and Manufacturing*, vol. 57, pp. 95–107, 2014, ISSN: 1359835X.

[10] D. Durville, I. Baydoun, H. Moustacas, G. Périé, and Y. Wielhorski, “Determining the initial configuration and characterizing the mechanical properties of 3D angle-interlock fabrics using finite element simulation,” *International Journal of Solids and Structures*, vol. 154, pp. 97–103, Dec. 2018, ISSN: 0020-7683.

[11] Z. Yang et al., “Modeling of 3D woven fibre structures by numerical simulation of the weaving process,” *Composites Science and Technology*, vol. 206, p. 108 679, Apr. 2021, ISSN: 0266-3538.

[12] T. Shao, S. Zhang, W. Liu, R. Liu, and W. Xu, “High-fidelity architecture modeling and compressive strength prediction of 3D woven composite material,” *Composite Structures*, vol. 354, p. 118 775, Jan. 2025, ISSN: 0263-8223.

[13] N. Naouar, E. Vidal-Salle, J. Schneider, E. Maire, and P. Boisse, “3D composite reinforcement meso F.E. analyses based on x-ray computed tomography,” *Composite Structures*, vol. 132, pp. 1094–1104, Nov. 2015, ISSN: 0263-8223.

[14] R. M. Auenhammer, L. P. Mikkelsen, L. E. Asp, and B. J. Blinzler, “Automated X-ray computer tomography segmentation method for finite element analysis of non-crimp fabric reinforced composites,” *Composite Structures*, vol. 256, no. August 2020, p. 113 136, 2021, ISSN: 02638223.

[15] Y. Sinchuk et al., “X-ray CT based multi-layer unit cell modeling of carbon fiber-reinforced textile composites: Segmentation, meshing and elastic property homogenization,” *Composite Structures*, vol. 298, p. 116 003, Oct. 2022, ISSN: 0263-8223.

[16] A. Ewert, B. Drach, K. Vasylevskyi, and I. Tsukrov, “Predicting the overall response of an orthogonal 3D woven composite using simulated and tomography-derived geometry,” *Composite Structures*, vol. 243, p. 112 169, Jul. 2020, ISSN: 0263-8223.

[17] Y. Pannier, P. Coupé, T. Garrigues, M. Gueguen, and P. Carré, “Automatic segmentation and fibre orientation estimation from low resolution X-ray computed tomography images of 3D woven composites,” *Composite Structures*, vol. 318, no. April, p. 117 087, 2023, ISSN: 02638223.

[18] C. Tang, J. Zou, Y. Xiong, B. Liang, and W. Zhang, “Automatic reconstruction of closely packed fabric composite RVEs using yarn-level micro-CT images processed by convolutional neural networks (CNNs) and based on physical characteristics,” *Composites Science and Technology*, vol. 252, no. April, p. 110 616, 2024, ISSN: 02663538.

[19] Z. Ying et al., “Geometric reconstruction of mesoscopic models for three-dimensional woven composites using micro-ct technology,” *Composite Structures*, vol. 371, p. 119 466, Nov. 2025, ISSN: 0263-8223.

[20] M. A. Ali, Q. Guan, R. Umer, W. J. Cantwell, and T. J. Zhang, “Deep learning based semantic segmentation of µCT images for creating digital material twins of fibrous reinforcements,” *Composites Part A: Applied Science and Manufacturing*, vol. 139, no. August, p. 106 131, 2020, ISSN: 1359835X.

[21] A. Tsamos, S. Evsevleev, R. Fioresi, F. Faglioni, and G. Bruno, “Synthetic Data Generation for Automatic Segmentation of X-ray Computed Tomography Reconstructions of Complex Microstructures,” *Journal of Imaging*, vol. 9, no. 2, 2023, ISSN: 2313433X.

[22] D. Horváth, G. Erdős, Z. Istenes, T. Horváth, and S. Földi, “Object detection using sim2real domain randomization for robotic applications,” *IEEE Transactions on Robotics*, vol. 39, no. 2, pp. 1225–1243, Apr. 2023, ISSN: 1941-0468.

[23] L. A. Feldkamp, L. C. Davis, and J. W. Kress, “Practical cone-beam algorithm,” *Journal of the Optical Society of America A*, vol. 1, no. 6, pp. 612–619, 1984.

[24] Z. Chen, V. D. Calhoun, and S. Chang, “Compensating the intensity fall-off effect in cone-beam tomography by an empirical weight formula,” *Applied Optics*, vol. 47, no. 32, pp. 6033–6039, 2008, ISSN: 15394522.

[25] N. Otsu, "A threshold selection method from gray-level histograms," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 9, no. 1, pp. 62–66, Jan. 1979, ISSN: 2168-2909.

[26] E. R. Dougherty, *Mathematical morphology in image processing* (Optical Science and Engineering). Boca Raton, FL: CRC Press, Sep. 1992.

[27] S. U. Lee, S. Chung Seok Yoon, and R. H. Park, "A comparative performance study of several global thresholding techniques for segmentation," *Computer Vision, Graphics and Image Processing*, vol. 52, no. 2, pp. 171–190, 1990, ISSN: 0734189X.

[28] J. B. Roerdink and A. Meijster, "The watershed transform: Definitions, algorithms and parallelization strategies," *Fundamenta Informaticae*, vol. 41, no. 1, 2, pp. 187–228, 2000, ISSN: 0169-2968.

[29] N. Naouar, E. Vidal-Sallé, J. Schneider, É. Maire, and P. Boisse, "Meso-scale FE analyses of textile composite reinforcement deformation based on X-ray computed tomography," *Composite Structures*, vol. 116, no. 1, pp. 165–176, 2014, ISSN: 02638223.

[30] O. Ronneberger, P. Fischer, and T. Brox, *U-net: Convolutional networks for biomedical image segmentation*, 2015.

[31] Z. Gao, L. Chen, S. Zhao, and Z. Zhang, "A deformation procedure using position-based dynamics to optimize the geometric model of woven fabrics," *Applied Mathematical Modelling*, vol. 137, p. 115 650, Jan. 2025, ISSN: 0307-904X.

[32] A. H. Chen, Z. Liu, Y. Yin, and C. Yuksel, "Vertex block descent," *ACM Transactions on Graphics (TOG)*, no. 116, 2024.

[33] C. Giles, E. Diaz, and C. Yuksel, "Augmented vertex block descent," *ACM Transactions on Graphics (Proceedings of SIGGRAPH 2025)*, vol. 44, no. 4, Aug. 2025, ISSN: 0730-0301.

Part II

Papers

PAPER A

Synthetic, automatically labelled training data for machine learning based X-ray CT image segmentation: Application to 3D-textile carbon fibre reinforced composites

Johan Friemann, Lars P. Mikkelsen, Carolyn Oddy, Martin Fagerström

Published in Composites Part B: Engineering,
Vol. 305, 112656, Oct. 2025.

https://doi.org/10.1016/j.compositesb.2025.112656

Composites Part B 305 (2025) 112656

Contents lists available at ScienceDirect

Composites Part B

journal homepage: www.elsevier.com/locate/composites

Synthetic, automatically labelled training data for machine learning based X-ray CT image segmentation: Application to 3D-textile carbon fibre reinforced composites

Johan Friemann [a],*, Lars P. Mikkelsen [b], Carolyn Oddy [a,c], Martin Fagerström [a]

[a] Material and Computational Mechanics, Department of Industrial and Materials Science, Chalmers University of Technology, SE-41296 Göteborg, Sweden
[b] Composites Manufacturing and Testing, Department of Wind and Energy Systems, Technical University of Denmark, DK-4000 Roskilde, Denmark
[c] GKN Aerospace Sweden, SE-461 81 Trollhättan, Sweden

ARTICLE INFO

Keywords:
Segmentation
X-ray CT
3D-textile reinforced composites
Machine learning
Open source software

ABSTRACT

Composite parts with 3D-textile reinforcement show promise in high-performance applications. For widespread use, accurate material characterisations are required. Characterisation of the textile architecture in the as-manufactured state may be performed with X-ray CT. Due to the similarity between the chemical composition of carbon fibres and epoxy based matrices, the contrast of X-ray CT scans is poor. Therefore, segmentation with classical methods is difficult or even impossible. Alternatively, machine learning based segmentation approaches may be used. One drawback of machine learning-based algorithms is the need for large datasets whose ground truth labellings require extensive manual labour. This can be circumvented by utilising automatically labelled synthetic X-ray CT data. In this work, a novel pipeline that generates synthetic CT image datasets, with automatically labelled ground truths, is developed. The pipeline is entirely based on free and/or open source software. It is demonstrated that segmentation model, trained on only such data, is able to accurately segment a real X-ray CT scan of a 3D-reinforced carbon fibre composite sample. A pixel-wise agreement of 88% is reached when compared to a manual segmentation. This implies potentially large time savings in segmentation tasks, which could accelerate characterisation of textile composites in their as-manufactured state.

1. Introduction

Composites with 3D-textile reinforcements have gained increased interest in high performance applications, owing to their promising out-of-plane mechanical properties per unit mass [1]. The out of plane reinforcing yarns make 3D-textile composites very resistant to delamination. For example, by distributing only 1% of the reinforcement yarns through the thickness in a 3D-woven carbon fibre reinforced polymer (CFRP), it is possible to increase the mode I interlaminar toughness by more than 10 fold when compared to a traditional laminated counterpart [2]. A noticeable improvement has also similarly been demonstrated with 3D-woven glass fibre composites [3]. It has also been demonstrated that 3D-woven composites possess significantly higher levels of damage resistance during impact than comparable 2D woven composites [4,5].

In order to enable the utilisation of 3D-textile composites in high performance engineering applications such as those in the aerospace industry, full material characterisation must be carried out. Generally speaking, 3D-textile composites are highly anisotropic and therefore require expensive and time consuming experimental testing campaigns to characterise the material. Any change in fibre or yarn type, yarn architecture or matrix material would then invalidate all previous tests and require a new testing campaign. One approach considered to circumvent this challenge is the use of representative volume element (RVEs) modelling of the mesostructure. These explicitly model the yarn architecture and the matrix that surrounds them. These models can be used to predict the homogenised properties of the material. As an added benefit, these models can also make it possible to study the micromechanical behaviour internally in the material and how the weaving architecture can affect the response. However, one issue that this method faces is the need for accurate geometries that describe the meso-scale features with sufficient and accurate detail.

To create such meso-scale geometrical representations, the simplest method is a direct geometric description of the woven architecture. Geometric modelling software such as TexGen [6] or WiseTex [7] use interpolating curves with associated cross sections to describe the weave, which can later be used to create a finite element mesh.

---

* Corresponding author.
E-mail address: johan.friemann@chalmers.se (J. Friemann).

https://doi.org/10.1016/j.compositesb.2025.112656
Received 14 April 2025; Received in revised form 16 May 2025; Accepted 27 May 2025

Alternatively, the geometry can be derived from simulating the weaving process itself. This is often done through the Digital Element Method (DEM), introduced by Wang and Sun [8] and later refined by Zhou et al. [9]. In DEM, yarns are considered as ensembles of chains of rod elements, between which contact conditions are considered. With this set-up, the weaving process can be simulated and realistic weaving geometries obtained. Both aforementioned methods can be combined by using the geometric description as a pre-processor to define the yarn pattern, followed by a DEM simulation to more accurately describe the material as-woven after compaction. This was proposed by Mahadik and Hallet [10], then later refined by Green et al. [11] and El Said et al. [12].

However, there are some drawbacks with these approaches. Geometric modelling is typically overly simplistic and may generate penetrating yarns, which is unacceptable for finite element meshes. Digital element simulations on the other hand are rich in contacts and are quite computationally expensive. By bearing the already expensive finite element analysis of the RVE in mind it can become prohibitively time consuming to generate geometries with DEM if many RVE configurations are to be investigated.

As an alternative, there is growing use of X-ray Computed Tomography (XRCT) to generate RVE geometries that are inferred from 3D reconstructions of scans of actual material samples [13–15]. Ewert et al. [16] showed that a model built on geometry derived from XRCT data outperformed models based on DEM derived geometry, yielding better estimates of constituent volume fractions and mechanical properties. However, to create a computational model from XRCT, each voxel of the reconstruction needs to be assigned its corresponding material such as matrix, yarn, or defects like voids (air pockets inside the matrix). Yarns also need their local fibre orientation and volume fractions specified. Assigning the voxels to their respective constituent phases is an image segmentation problem. Segmenting XRCT images can be challenging, especially for woven CFRPs where the contrast between matrix and matrix infused carbon fibre yarns is particularly low due to their similar chemical composition. Therefore, noise sensitive classical segmentation methods, like Otsu thresholding and watershed, often fail, requiring more sophisticated methods [17,18]. Naouar et al. [19] have successfully used structure tensor analysis in combination with classical image processing methods to segment CT scans of 2D CFRP weaves, and texture analysis combined with classical image processing to segment 3D-woven glass fibre scans [20]. Fourrier et al. [21] also demonstrated how structure tensor based segmentation works for low contrast scans. Albeit able to outperform classic methods, their method is dependent on structural data being distinguishable in the XRCT images. This may not be the case for a large field of view scan which may be needed given that the side length of a woven composite unit cell can be upwards of several cm. Through another alternative method Bénézech and Couégnat [22], Sinchuk et al. [14], and Pidou-Brion and Guilloux [23] successfully utilised variational approaches, where an energy like functional describing the difference between a segmentation and XRCT image is minimised. However, while comparing their results Sinchuk et al. [14] demonstrated how a machine learning based model outperformed their variational method.

Machine learning, specifically deep learning methods, seem to dominate contemporary segmentation of 3D-woven CFRP scans and show great promise. Some examples include Pannier et al. [24] that segmented low resolution scans of 3D-woven composites with a neural net applied to structure tensor features, and Tang et al. [15] that segmented closely packed fabric composites with convolutional neural networks. However, the methods do in principle require manual segmentation of the training data ground truths. This requires a painstaking amount of labour for a large dataset. Blusseau et al. [25] reduced the manual labelling workload by combining manually labelled yarn centres with "pseudo labels" derived from classical image processing methods for yarn cross sections. A different method was investigated by Mendoza et al. [26], where mechanical simulations based on an XRCT volume were used to generate a labelled dataset. This dataset was then used to train a neural network to generate "pseudo CT" images using the original image as the ground truth. A second network was then trained using the pseudo images as input and the original segmented images as ground truths. Finally this second network was used to run inference on real XRCT data. Zheng et al. [27] used another different approach where they fortified their training dataset with artificial XRCT images based on statistics of their original dataset.

One unifying factor in all the machine learning based methods presented above is their high degree of specialisation. The data, regardless if it is automatically or manually labelled, is limited to one sample and/or one XRCT scanner set-up. A completely different method for generating training data is to utilise solely synthetic data, for example Ali et al. [28] and Konopczynsky [29]. Ali et al. specifically, have trained a segmentation algorithm on heuristic synthetic tomography data successfully. However, one promising opening for improvement would be to demonstrate inference on real XRCT data or on different yarn architectures. In a later publication Ali et al. [30] improve their methodology and create a model that shows impressive generalisation, but they abandon synthetic training data and return to manual labelling. Outside the field of fibre reinforced polymers, synthetic data has successfully been used to train segmentation models by several authors. Tsamos et al. [31] utilised a fully automatically generated dataset for a six-phase Al–Si alloy composite reinforced with ceramic fibres and particles, but did not base the synthetic data on physical first principles. In contrast, Fok et al. [32], used physically simulated synthetic X-ray images (not CT) for segmenting clavicles in chest X-ray images. Their synthetic X-ray projections were however based on tomographic volumes collected from real patients, and did further not utilise a fully synthetic dataset.

The purpose of this work is to investigate the potential of using fully synthetic XRCT data to train a machine learning based segmentation algorithm aimed for segmentation of physical scans of 3D-woven CFRP samples. For this purpose, a novel pipeline for synthetic data generation is presented. For each virtual (or synthetic) material sample, the pipeline generates the weave geometry, simulates an XRCT scan based on first principles, reconstructs the X-ray projections obtained from the simulated XRCT scan, and automatically labels the reconstructed voxels.

By performing a physically-based simulation and reconstruction procedure, it is demonstrated that each synthetic data point (i.e. fully labelled virtual scan) includes the type of noise and reconstruction artifacts present in real XRCT data. Additionally, another benefit from this pipeline is that the segmented ground truth can be derived automatically from the simulation input in a straightforward fashion. Furthermore, by making the data generation pipeline computationally efficient, it is possible to generate a large and varied dataset. This dataset can then be used to train a machine learning model aimed to segment XRCT images of a variety of weaves scanned on a variety of equipment. It is demonstrated that a machine learning algorithm trained on such a dataset reaches a pixel-wise accuracy of 88% when compared to a similar segmentation performed by a human.

The pipeline is completely based on commonly available free and/or open source software. The entire code repository with documentation is openly available online [33]. An XRCT scan of a method validation sample, and a corresponding manually performed segmentation, is hosted freely online [34], and an additional testing sample is hosted at [35]. Due to the large size of the generated synthetic dataset (roughly 700 GB), it is not hosted. An equivalent dataset can be generated by using the provided code and parameters. However, if the exact dataset is requested it will be made available. To the authors' best knowledge no other work has presented a similar strategy for segmenting XRCT images, neither for 3D-woven composites nor other materials research.

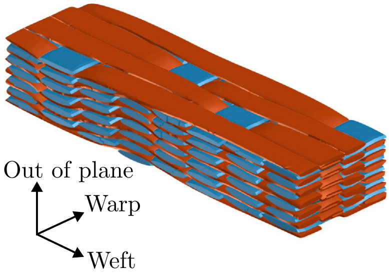


Fig. 1. The studied layer-to-layer angle interlock material's unit cell is shown. Weft yarns are shown in red, and warp yarns are shown in blue. Original graphics provided by courtesy of Oddy [41]. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

2. Method

The investigation of whether a synthetic XRCT dataset can be used to train an algorithm that can generalise to real data is based around a specific 3D-woven CFRP sample. The sample is a layer to layer, angle interlock CFRP. A schematic of the unit cell is displayed in Fig. 1. The weft yarns consist of 12k count IM7 fibres, the warp yarns are 24k count IM7 fibres. The IM7 fibres [36] have a diameter of 5.2 µm and a density of 1.78 g cm^{-3}.

The dry textile preform is injected with RTM6 epoxy [37] using resin transfer moulding (RTM). This epoxy system is a two part epoxy with a cured density of 1.14 g cm^{-3}. The exact chemical profile of RTM6 is not made publicly available by the supplier. It is however important when estimating the material attenuation. Both [38,39] make reference to the chemical composition of RTM6. However, through discussion with Westman [40] it is believed that there may be an error in these previous studies, whereby a slightly modified (corrected) chemical composition is assumed in the following.

For the purpose of this study it is assumed that part A consists of Tetraglycidyl Methylene Dianiline, while part B is made up of a combination of 4,4'-methylenebis(2,6-dimethylanline) and 4,4'-methylenebis (2isopropyl,6-methylanline). The ratios of constituents in part B is not disclosed due to commercial interests. Thereby, in this work they are assumed to be present in the same amount. The weight ratio of A:B is given in the RTM6 data sheet as 100:68.1. The chemical formulae for the three constituents are respectively $C_{25}H_{30}N_2O_4$, $C_{17}H_{22}N_2$, and $C_{21}H_{30}N_2$. It should be mentioned that the composite also includes a few glass fibre tracer yarns that were used as aids during manufacturing. These will cause artifacts in any XRCT scans, but are not considered when generating the synthetic datasets.

The unit cell is quite large compared to the fibre diameter being 21.5 mm long in the weft direction, 6.5 mm wide in the warp direction, and 4 mm thick in the out of plane direction. The sample is cut out from a cured sheet into 3 pieces of roughly 2 by 2 unit cells that are stacked on top of each other in the out of plane direction. See Fig. 2 for an image of the sample. Stacking is performed in order to get a good aspect ratio, enabling uniform attenuation during the scan. The size of the sample is motivated by the desire to produce XRCT data with a large number unit cells. If an algorithm is able to successfully segment many unit cells at once, it would provide a user with the ability to perform valuable statistical analysis on varied RVE geometries.

In the first part of this section, an experimental XRCT scan of the sample considered for model testing is described. The second part and main contribution of this work consists of a detailed description of the proposed data generation pipeline. The section concludes with a description of the utilised machine learning algorithm and how it is trained and tested/validated. The focus of this work is the demonstration of the viability of synthetic XRCT data for training segmentation algorithms. As a consequence the machine learning procedure for automatic segmentation has lower priority.

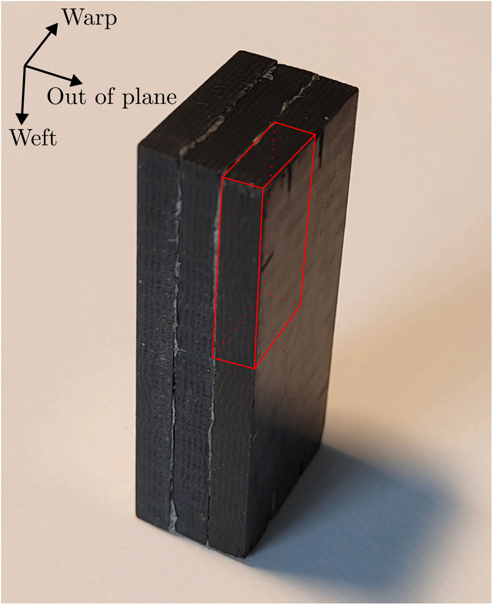


Fig. 2. The sample used in this study is displayed. The red box has been added on the image to visualise the approximate size of one unit cell in the sample, cf. the cell shown in Fig. 1. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

2.1. Scan and reconstruction of a real sample for testing

A *Zeiss XRadia 410 Versa* lab based XRCT scanner is used to scan the described sample. The scanner uses a *Hamamatsu Photonics L8121-03* micro focus X-ray source with a tungsten target. X-ray filters can be inserted at the source aperture, but in this work no filter is used. There is however most likely an undocumented X-ray filter somewhere on the X-ray source side of the scanner. It is assumed to be 1 mm of aluminium. The X-ray source documentation only mentions a 200 µm beryllium filter, but initial investigations show agreement between an attenuation model and experiment if a 1 mm aluminium filter is used. Furthermore, there is no mention on the angle of the X-ray source anode, and it is assumed to be 12° as it is the most common value. The scanner has an *ANDOR iKon-L DW936N-BV-556* CCD detector. The detector array consists of 2 048 by 2 048 pixels with dimensions 13.5 µm by 13.5 µm. The XRadia 410 Versa scanner is equipped with an optical system between the X-ray scintillation panel and the CCD detector. This system enables higher magnification, or a larger field of view than a pure detector panel set-up.

The sample is placed 80 mm away from the X-ray source, and 150 mm away from the detector. The Large Field of View (LFOV) detector objective with 0.4 times magnification is used. The X-ray tube voltage is set to 40 kV and the tube power is set to 10 W. The exposure time is set to 5 s. No additional X-ray filter is installed. The tomographic scan is performed by rotating the sample 360° in increments and saving the projections at each step. The scan is performed with 3001 projections and binning 4. Binning n refers to an averaging of the intensity of n by n pixel values. This results in an image n^2 times smaller than the original. The combined geometric and objective magnification, together with the binning results in an effective voxel size of 46.96 µm. Refer to Table 1 for a summary of the used measurement settings.

Table 1
The equipment settings that were used for the tomographic scan.

Parameter	Value	Unit
Source to object distance	80	mm
Object to detector distance	150	mm
Objective	LFOV	–
Binning	4	–
Source tube voltage	40	kV
Source tube power	10	W
Exposure time	5	s
Scanning angle	360	°
Number of projections	3001	–

2.1.1. Tomographic reconstruction

As the scan in this work is done with a cone beam it is reconstructed with the Feldkamp–Davis–Kress (FDK) technique [42], which is a generalisation of filtered back projection. This procedure is not an exact reconstruction algorithm (except the central vertical axis slice) unlike filtered back projection for a parallel beam set-up. The approximations in standard implementations of FDK typically result in a longitudinal fall-off effect where densities are underestimated at the extreme ends of the vertical axis [43]. In this work, the stack of projections is reconstructed with help of the ASTRA toolbox [44] that has a convenient python interface. The ASTRA toolbox has a GPU implementation of FDK that allows for rapid highly parallelised reconstruction.

2.1.2. Manual segmentation

In order to evaluate the feasibility of fully synthetic training data for segmenting experimentally derived tomographic reconstructions, the reconstruction of the scan is segmented by hand. The hand segmentation is performed with the open source software 3D Slicer [45]. It is emphasised that this manually segmented data is not seen by a model during training, but is only used for testing the model against an independent ground truth. Owing to poor contrast and human factors, it cannot be claimed that a manual segmentation represents a perfect ground truth segmentation. Therefore, any performance metrics cannot be taken as a direct measure of accuracy, and should therefore instead be interpreted as similarity to an alternative good faith segmentation effort.

2.2. Estimation of scanner blurring

For the purpose of investigating whether blurring due to the optical system or sensor imperfection needs to be accounted for in the generation of synthetic tomography images, the Point Spread Function (PSF) of the scanner's detector and optical system is estimated. It is assumed that a pixel hit by an X-ray affects its neighbouring pixels in a rotational symmetric way. Under that assumption the PSF can be applied through two subsequent applications of the line spread function orthogonal to each other. The line spread function is found through the derivative of the edge spread function, which in turn can be measured through the X-ray imaging of a sharp edge made out of a material impassible to X-rays.

To measure the edge spread function, an adjustable frame containing two sets of orthogonal sharp edges made out of tungsten is placed in the scanner sample position. The detector is placed as close as possible to the edge, and the source is placed far away. This is done in order to achieve a large effective pixel size, thereby minimising the influence of blurring due to finite source spot size. The detector is indicated by the XRCT scanner as 32 mm away from the sample centre and the source is 200 mm away. This results in an effective pixel size of 29 μm. The LFOV (0.4× magnification) objective was used. However, a lack of appropriate mounting brackets entailed placing the frame on the sample platform manually, resulting in the edge not coinciding with the sample position as indicated by the XRCT scanner. The error introduced by this inaccuracy, which is on the order of a few millimetres, is deemed to not affect the outcome of the experiment. Since the effective pixel size is so large, the resulting error in the effective voxel size is around 0.5% per mm.[1]

Table 2
The equipment settings that were used for capturing a projection estimating the edge spread function.

Parameter	Value	Unit
Source to object distance	200	mm
Object to detector distance	32	mm
Objective	LFOV	–
Binning	1	–
Source tube voltage	40	kV
Source tube power	10	W
Exposure time	10	s

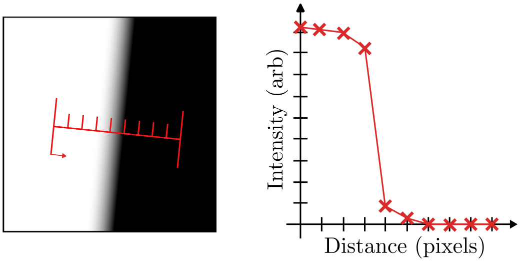


Fig. 3. An artificial edge spread measurement is shown. The intensity as a function of pixel distance (synthetic data for demonstration) is shown next to an artificial X-ray projection of a dense metal edge. The red ruler, given in pixel units, is orthogonal to the edge. The average intensities over bands parallel to the edge is sampled along the ruler. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

The X-ray source potential is set to 40 kV, and the power is set to 10 W. Binning 1 (no binning) and an exposure time of 10 s is utilised. An X-ray projection is taken, and the pixel intensity values are sampled along a line orthogonal to the edge. Local fluctuations are reduced by letting the line be 500 pixels wide (parallel to the edge) and computing the average intensity along the width. See Fig. 3 for a visualisation example. A summary of the scanner settings can be found in Table 2. The PSF is estimated from the derivative of this intensity. As the values correspond to discrete pixel coordinates the derivative is computed as the differences between subsequent intensities through

$$\frac{dI}{dx} = I_{n+1} - I_n, \tag{1}$$

where the differential is estimated as $dx = 1$, since the intensities are sampled pixel-wise.

2.3. Generation of synthetic training data

Training data is generated through a three step pipeline. In the first step, surface triangle mesh data that represents the weft, warp, and matrix geometry is generated. The second step entails performing a ray-casting simulation of X-ray attenuation through the meshes, and reconstructing the projection data into a 3D XRCT volume. The final step consists of a solid voxelisation of the meshes in the reference frame of the reconstruction. This represents a one to one perfect segmentation of the XRCT volume for the given scanning resolution.

In essence, a XRCT tomographic reconstruction (input) and its corresponding segmentation (output/ground truth) then constitutes a data point for training a segmentation algorithm. However in practice,

---

[1] $13.5\,\mu\text{m} \cdot (200\,\text{mm} \pm \delta\,\text{mm})/(200\,\text{mm} + 32\,\text{mm})/0.4 \approx 29\,\mu\text{m} \pm 0.15\,\mu\text{m} \cdot \delta$, for some perturbation δ.

due to large memory requirements for processing entire volumes, a data point is made up of one slice from the reconstruction (input) together with the corresponding slice from the segmentation (output/ground truth).

The proposed pipeline is implemented in Python and allows for high control of parameters and can be used to create a varied dataset with a high degree of domain randomisation, thus enabling the training of a robust algorithm. Domain randomisation entails handling the fact that the model is inexact, or the exact model parameters are unknown. If parameters are randomly sampled from a large enough domain, the so-called reality gap can be shrunk [46]. In practice, this means that for example, although the number of filaments within a yarn are well known and controlled precisely by the supplier, the parameter set used to calculate the attenuation properties will be sampled from a range of possible values. The same can be said for fibre density, fibre diameter etc. All of the outlined steps of the proposed pipeline are described in detail in the following sections.

2.3.1. Generation of woven geometry

The first step of the pipeline is generating closed surface meshes that are representative of the studied composite sample. Woven geometries are generated with the aforementioned TexGen software. TexGen describes yarns as interpolated space curves with associated cross sections at the interpolation nodes. In this work the python API to TexGen is utilised. Geometries that resemble the described woven sample are generated by combining the pre-existing 3D weave generation functionality with the layered textile functionality. The textile unit cell is constructed by defining the unit cell dimensions, the number of yarns, the yarn pattern (cross over points of weft and warp yarns), and the yarn dimensions and spacing.

Out of the box, TexGen generates unit cells that qualitatively replicates the ones in the physical sample. However, it can be difficult to control the waviness in the yarns as it is directly determined by TexGen from the unit cell size and yarn dimensions. In the physical 3D woven sample, the weft yarns display a higher degree of waviness than the warp yarns, cf. the fairly straight weft yarns in Fig. 1. In contrast, a TexGen 3D woven unit cell with the same specifications display the opposite, showing high waviness in the warp yarns. This is counteracted by also generating unit cells with a reversed aspect ratio, yielding samples with high weft crimp.

To fully replicate the scanned sample, a repeated number of unit cells with variations need to be generated. One slightly unorthodox TexGen application in the pipeline is letting a TexGen unit cell consist of many unit cells through a tiling of the original unit cell. Random perturbation of the interpolation nodes and random scaling and rotation of the cross sections allows for a continuous representation of multiple unit cells displaying variations as expected in a real sample.

The pipeline saves the weft yarns, warp yarns, and matrix as separate mesh files, something that is needed for applying different properties to the different phases in the subsequent XRCT simulation. It is worth noting that the matrix mesh is simply the smallest rectangular cuboid containing all the yarns making up the sample. It has 2 triangles per face, resulting in 12 triangles. TexGen exports triangular surface meshes with the normal pointing outwards.

TexGen has a tendency to create overlapping yarn geometries for high volume fraction weaves. In this work this is handled by a mesh post processing step. The software PyMeshLab, a Python API to the MeshLab software [47], is used to perform a Boolean difference operation as described by Zhou et al. [48]. Non-penetration is guaranteed by cutting one type of yarn mesh (weft or warp) out of the other. The meshes are also transformed with PyMeshLab such that their coordinate origins lies in the mesh geometry centre of mass.

2.3.2. Simulation of computed tomography

As a second stage in the pipeline, a virtual X-ray tomography scan of the meshes generated in the previous step is simulated, and a subsequent reconstruction is performed. Simulation of X-ray projections are performed with the help of the software gVirtualXray [49,50]. Fundamentally, it is based on an algorithm by Freud et al. [51] but has been is implemented to take advantage of the highly parallelised real time capabilities of the rendering pipeline that GPU's are designed on by using OpenGL. The algorithm in gVirtualXray is based on the Beer–Lambert attenuation law, and can model the attenuation of polychromatic X-rays. It has been shown by Pointon et al. that gVirtualXray creates accurate X-ray projections that are comparable in quality to more sophisticated (and expensive) X-ray physics models [52].

The materials of the phases are defined by listing the phases' constituent atomic numbers, their ratios, and the density of the materials. This will affect the attenuation coefficients as a function of photon energy. The library xraylib [53] is used by gVirtualXray to fetch the photon cross sections used to compute the attenuation coefficients. The glass fibre tracer yarns are not taken into account in this work. It is also important to state that the yarns in a 3D-woven CFRP are not homogeneous. The attenuation properties of the yarns must therefore be approximated by homogenisation before the simulation. For this purpose, the rule of mixtures is used. First, the matrix material's atomic weight ratios are estimated by weighting the constituent atomic mass ratios by their weight fractions. The effective yarn atomic mass ratios and yarn densities are estimated by a rule of mixtures using the weight fractions of fibre and matrix material inside the two yarn types. The fibre volume fraction is determined by dividing an estimated yarn cross sectional area in an XRCT slice with the yarn fibre counts multiplied with the fibre area. It is assumed as constant per yarn type which is a simplification, but the yarns show fairly small variation in cross sectional area in the real sample. The weight fractions of matrix and fibre in the yarns then follows by multiplication of their respective densities.

It should be pointed out that the software does not account for secondary photons reaching the detector through scattering. For X-ray photons emitted from a 40 kV tube potential, tungsten target, X-ray source filtered by 1 mm of aluminium the majority of photons will have an energy around 30 keV. For carbon based materials (low atomic number) the dominating attenuation effect at these energies will be photoelectric absorption of photons. Since Compton and Rayleigh scattering are the least common attenuation modes at these energies [54] it is deemed in this work that neglecting scattering will not significantly impact the realism of the simulation.

The pipeline utilises the python API to gVirtualXray for ease of implementation. All aspects of the simulations are configurable through this API, and can be perturbed for domain randomisation. The source is configured by setting the X-ray tube voltage, tube power, anode angle, exposure time, and eventual X-ray filter material and thickness. In the current work the X-ray spectrum that is input to gVirtualXray is generated based on the source parameters with the software SpekPy that shows good agreement with established X-ray physics simulators [55].

Each virtual sample is set up by loading the meshes of the individual phases and placing them together in relation to the source and detector. The relative distances will determine the geometric magnification of X-ray projections. The sample can furthermore be tilted and offset through a series of rigid body transformations to more accurately match an imperfectly set up scan environment. It is worth noting that the mesh of the matrix phase of a sample envelops the other phases. This is handled by gVirtualXray by setting the matrix's mesh as an "outer" mesh while loading it. This will subtract the matrix materials contribution to the attenuation for the distance traversed through the enveloped meshes, preventing counting attenuation for both the matrix and the "inner" mesh's material.

The detector is configurable by setting the pixel size, number of pixels, and binning. The detector can be configured to be in photon counting mode, or energy integrating mode. Here energy integrating mode is chosen to be comparable with the equipment in Section 2.1. In this work, all photons that reach the detector is assumed to deposit their entire energy. The energy efficiency of the scintillator material and the CCD is treated as perfect. This assumption is made out of necessity, since no information on the scintillator used by the mentioned equipment is available. The simulation software does not have the ability to model optics between the detector scintillator and CCD. In the pipeline the optics are therefore handled by simply scaling the detector pixels by the inverse of the magnification. Similarly, detector pixel binning is handled by scaling the pixel area by the binning number squared. These two adjustment will have the same effect on the reconstruction voxel sizes as the physical set up, but will not capture eventual blurring added by the optics. Blurring therefore needs to be added through post processing. Finally, the user chooses the scanning angle and the number of X-ray projections to capture.

After setting up the sample the tomographic simulation is performed by using the mentioned algorithm to capture projections and rotating the sample by the scanning angle divided by the number of projections between each projection. Blurring due to finite X-ray source finite focal spot size, detector optics, and CCD leakage is added in a post processing step by convolving the projections with a Point Spread Function (PSF). The PSF is approximated with a Gaussian kernel. Here, it is assumed that the blurring is independent of geometric magnification.

Furthermore, focal spot blur is not the dominating source of blurring for the resolutions modelled here. Focal spots achieved in scans that are comparable to the one described in Section 2.1 is on the scale of a few micrometers. This will result in a geometric blur smaller than the detector pixels for the levels of magnification studied here. Focal spot effects are therefore not explicitly modelled with gVirtualXray in this work due to the need for multi-sampled point sources which increases the computational time significantly.

Photon noise is an inevitable form of noise that arises due to the quantised nature of light, and is typically the dominating form of noise for high photon fluxes. Therefore, as an additional post processing step in the pipeline, Poisson distributed photon noise is added to each projection independently. The probability density function of a Poisson distributed variable is given by

$$f(k;\lambda) = \frac{\lambda^k \exp(-\lambda)}{k!}, \tag{2}$$

where k is the observed number of events and λ is the expected number of events. For photonic noise λ corresponds to the expected number of photons hitting a detector pixel. This depends on the photon flux, which in turn depends on the source flux as computed by SpekPy, the detector pixel size, and the exposure time. In this work other types of noise such as electronic noise in the detector are neglected.

Each X-ray projection is flat field corrected by scaling the detected integrated energy values by the expected energy per pixel that would be integrated by the detector if there was no sample. The flat field consists of the average of 30 images where each pixel is distributed according to Eq. (2), where the expected number of photons λ is determined from the spectrum simulator. The flat field correction is carried out as per standard for XRCT reconstruction. The choice of flat field sampling method is motivated by the desire to model noise features that are present in each projection, resulting in ring artifacts present in real CT-images.

The tomography simulation concludes with performing a tomographic reconstruction of the simulated projections. The ASTRA toolbox's implementation of the FDK algorithm is also used here. By using the same reconstruction method for the real XRCT data as for the synthetic a more fair comparison is enabled.

2.3.3. *Segmentation of simulated tomographic reconstructions*

The third and final stage of the pipeline consists of a segmentation of the meshed geometries generated in the first step. In order to create a training data point out of a tomographic image, the ground truth segmentation needs to be provided. Thereby, each voxel in the simulated reconstructed volume needs to be assigned its constituent class. Since the simulated scans are performed on surface meshes, a solid voxelisation of said surface meshes, obtained in the frame of reference of the reconstruction, would represent such a ground truth segmentation.

To obtain a solid voxelisation, the meshes are first transformed such that they are parametrised in the same frame of reference as used for the sample in the corresponding tomography simulation. That is, one voxel in the voxelisation corresponds to one voxel with the same size and location as a voxel in the tomographic reconstruction. Since the locations of the nodes of the surface meshes are defined in an axis aligned Cartesian coordinate system, with the origin placed in the sample centre of gravity, the transformation is easily achieved through a scaling, followed by a rigid body rotation, followed by a translation. A visualisation of the transformation process is found in Fig. 4. After transforming the meshes each material phase is voxelised separately. The individual constituent data is however stored in the same data structure, where each constituent is represented by a different integer value.

For the voxelisation, it is proposed to use a direct, triangle-parallel solid voxelisation algorithm, similar to the one presented by Schwartz and Seidel [56] but with some modifications to handle geometries that intersect the voxel grid bounding box (cf. box in Fig. 5) as well as and multiple material classes. Although Schwartz and Seidel present some more optimised methods in their work, the direct algorithm is chosen here for its simplicity and ease of implementation.[2] As such, the occupancy of each voxel is represented by a separate non-zero integer in an array whose length is equal to the number of voxels. The array is given row-major order where the x-coordinates correspond to the columns, the y-coordinates correspond to the rows, and the z-coordinates correspond to the slices.

At the start of the algorithm, this array is filled with zeros. Then, for each triangular surface element the following steps are carried out in a parallelised manner. Refer to Fig. 5 for a conceptual visualisation. First, the bounding box in the yz-plane of each triangular element is determined. Thereafter, the range of voxel centres in the yz-plane that are inside the bounding box is found. The xy-plane is an arbitrary choice, that does not affect the end result of the voxelisation. If the range of voxel candidates is non empty the triangle is projected along the x-axis on to the yz-plane. Each voxel centre in the determined range is then tested to see if they are inside or outside the triangle by utilising edge functions (analogous to triangle rasterisation) [57]. If a centre is inside the triangle, it is projected back along the x-axis on to the plane spanned by the triangle in 3D space. In a subsequent step, all the voxels whose centre has an x-coordinate greater than the projected coordinate is bit-flipped by performing an atomic XOR operation between the current array value and 1. This will make the voxel occupied if and only if it is currently empty. Therefore, it is emphasised that as the mesh is closed, erroneously flipped voxels will be flipped back when the triangle on the opposite end of the surface is treated, cf. Fig. 6.

The voxelisation procedure outlined above is performed on each individual phase mesh consecutively, utilising the same integer array for bit-flipping. The XOR operation is performed with different integers for each phase, that consist of only one non-zero bit (increasing powers of 2). Since the yarn meshes are guaranteed to be non overlapping, voxels that belong to one yarn type will always be flipped an even

---

[2] The potential gain in computational speed (in the order of 10 s of ms) is considered unnecessary given the growth in complexity and that the voxelisation step is not the bottleneck in the data-generation pipeline.

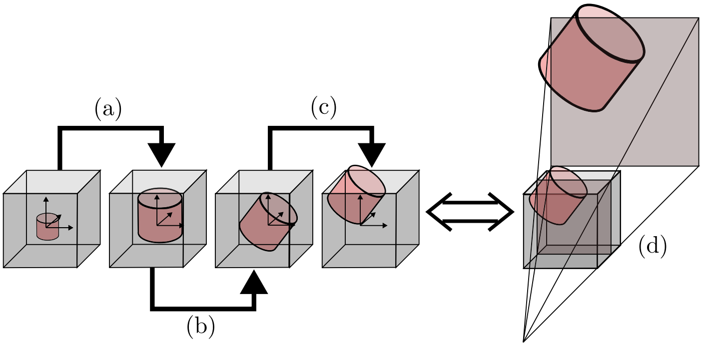


Fig. 4. The process of transforming a mesh into the tomographic reconstruction's frame of reference is shown. A scaling (a) is followed by a rigid body rotation (b), and a translation (c). Note that the resulting frame is equivalent to the frame of the tomographic reconstruction (d).

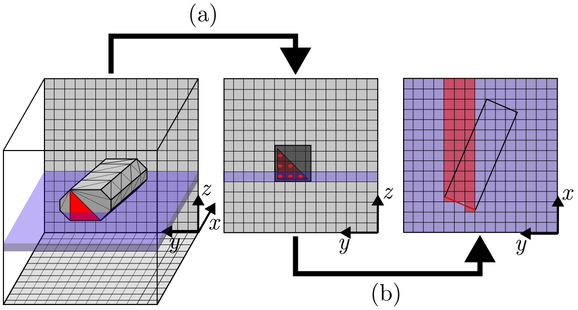


Fig. 5. The triangle parallel voxelisation is demonstrated for one triangle. Only contributions of the red triangle to voxels in the blue plane are visualised. First the triangle is projected on the yz-plane (a) and the bounding box is computed (dark box). Thereafter the voxel centres that are found to be inside the triangle are determined (red dots). The centres are projected back to the original 3D space (b) and any voxel in the corresponding row, with a larger x-coordinate is flipped, where the black rectangle indicates the geometry exterior in the plane. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

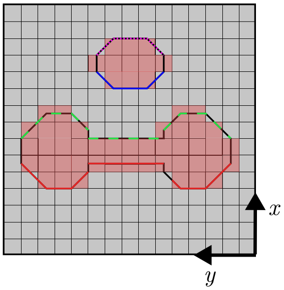


Fig. 6. A demonstration of repeated voxel flipping in one slice (fixed z). Voxels above the bottom shape, but outside the upper shape, will be flipped twice due to triangles along the red edge and green dashed edge, or four times due to the contribution of triangles on the blue edge and magenta dotted edge. These voxels will hence remain inactive. All other voxels will be flipped an odd number of times and will be set to active. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

number of times while being tested by the other yarn type, leaving them unchanged.

The exception is the matrix phase, which envelops all the yarn meshes. This phase is handled last and will inadvertently alter voxels belonging to the yarns. This is remedied by setting all but the least significant bit in each entry of the array to 0 after the matrix phase has been treated. As the matrix phase is treated last it must therefore correspond to the most significant bit, and will thus be set to 0 for voxels containing a bit with a smaller value. Finally, the most significant bit in the resulting integers is found, which results in an array with 0 representing air, 1 representing weft yarns, 2 representing warp yarns, and 3 representing matrix.

The algorithm is further extended in this work to support meshes that extend beyond the voxel grid, by augmenting the bounding box checking stage. This is important when treating a sample that intersects the bounds of the reconstruction volume, cf. the geometry sticking outside the final cube in 4. If a voxel candidate is outside the yz-bounds of the grid it is ignored. If a candidate after projection on the triangle is outside the upper x-bound it is also ignored. If the projection is below the lower x-bound all the voxels in that row is flipped. See Fig. 7 for an example. In summary: for each mesh, preprocess as shown in Fig. 4. Thereafter, for each triangle, carry out the basic operation shown in Fig. 5. Fig. 6 demonstrates how voxels flipped erroneously by one triangle are amended by subsequent triangles. Finally, Fig. 7 clarifies how the proposed algorithm has been modified to handle special cases not treated in the original algorithm.

In the pipeline, the proposed voxelisation scheme is implemented to run on the GPU through a custom CUDA kernel that is run through the package CuPy [58]. CuPy is an open source array library for interfacing with CUDA toolkit libraries for GPU computations.

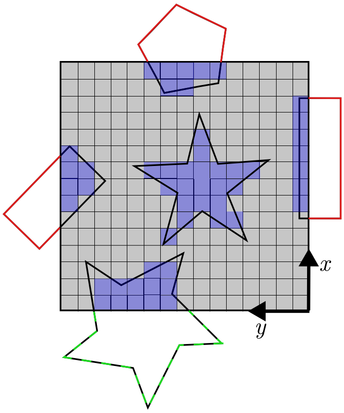


Fig. 7. The modified bounds checking is visualised in one slice (fixed z). Triangles along the red edges are ignored, while triangles along the green dashed edge flips all voxels in their corresponding rows. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

2.4. Choice and training of segmentation algorithm

The machine learning library PyTorch [59] is utilised. As the focus of this study is the viability of using synthetic training data, a custom designed machine learning algorithm is out of scope. Therefore, the prebuilt `fcn_resnet50` segmentation network directly available in the PyTorch API is used. It is a Fully Convolutional Network (FCN) segmentation head [60] with a 50 layer ResNet [61] backbone. The only adjustments to the network architecture is to the first and last layers. The first layer of the backbone is modified to account for single channel (gray scale) data instead of 3 channel (RGB). The final layer in the FCN segmentation head is modified to have 4 output classes (air, weft, warp, matrix) instead of the 20 classes in the prebuilt network. As the yarn segmentation is such a different task than normal RGB image segmentation (dogs, boats, humans, etc.) the network is initialised with random weights instead of using a pre-trained network. Out of the prebuilt segmentation models in PyTorch, the chosen model demonstrated the most promise during initial investigations.

It should be remarked that the adopted model is designed for segmentation of individual images. The tomography segmentation task is however intrinsically three dimensional. An efficient segmentation model should in theory benefit from exploiting the information available from the connectivity between tomographic slices. A 3D-aware model is considered out of scope in this work and tomographic volumes are segmented on a slice per slice basis.

A two stage training procedure is performed. In the first stage the segmentation model is trained with Stochastic Gradient Descent (SGD) with momentum and weight decay. A base learning rate of 0.001, a momentum of 0.99, and weight decay of 0.00001 is used. The batch size is 4.

The model is trained for 10 epochs, that is the entire training set is seen by the model 10 times during training. A cosine annealing learning rate scheduler is utilised. It entails letting the learning rate vary like a cosine function of the SGD iterations. Here, the period of the cosine function is chosen to be two training epochs. This results in an epoch of shrinking learning rate followed by an epoch of growing learning rate. Furthermore, a class-weighted cross-entropy loss function is utilised, where median frequency balancing is used for the class weights. The weights are computed as

$$w_i = \frac{\underset{j\in\{\text{classes}\}}{\text{median}}(n_j)}{n_i} \tag{3}$$

where n_i is the frequency of voxels in class i. Class balancing is used to prevent an algorithm from minimising its error by focusing on segmenting overrepresented classes and ignoring more rare classes. Finally, z-score normalisation is performed on the model input data, which is based on the statistics of the entire set of training data. Any inference performed with the model needs to be normalised in the same way. This means subtracting the mean and dividing by the standard deviation from all pixels in a slice to be segmented.

In the second training stage, the model is trained with identical hyper parameters for an additional epoch, but with an initial learning rate of 0.0001, and linearly decaying learning rate instead of annealing. That is, at each SGD iteration the learning rate, which starts at 0.0001, is reduced by a constant amount such that it is 0 at the last iteration. This additional stage is included to ensure that training ends when the learning rate is at a minimum.

3. Results

The results section begins by presenting the reconstruction of the XRCT scan of the real 3D-woven composite sample. Thereafter, the results of the line spread experiment is summarised. This information is used in the parameter choice for the generation of the synthetic datasets, which is presented next. Subsequently, the trained segmentation model is evaluated. The results conclude with a validation of the trained segmentation model on the real XRCT scan of the 3D-woven sample, and an additional XRCT scan of an extra testing sample that has a fundamentally different structure. The datasets were generated, and the segmentation model was trained on a laptop with an Intel i9-13900H CPU, 32 GB of RAM, a 1 TB NVMe M.2 Performance SSD, and an Nvidia RTX 4080 Laptop GPU (12 GB VRAM).

3.1. Real sample reconstruction

The tomographic scan took approximately 4 h. In Fig. 8, three representative slices of the reconstructed scan, along different axes, are shown. A histogram of the intensities in the entire reconstructed volume is later given in the solid plot in Fig. 13. The fact that the two yarn types do not give rise to distinct peaks demonstrates that a thresholding based segmentation is unsuitable for this reconstruction.

3.2. Manual segmentation of sample

First, using 3D Slicer everything inside the sample was marked as matrix, while everything outside was marked as air. Thereafter, the warp yarns were traced out every 10th slice in the plane normal to the warp cross sections. The warp yarns were then constructed by using the software's interpolation feature. After visual inspection of the resulting yarns, an additional smoothening step was added, with the purpose of obtaining a better match with reality. The smoothening was performed by first growing the segmentation by 1 voxel using the so-called margin function, whereafter a smoothing filter was applied. In the next step, the weft yarns were segmented in the nearly the same fashion, only excluding the margin operation.

The manual segmentation process took one person two working weeks to carry out. A similar 3D view and the same slices that are shown in Fig. 8 is shown for the manual segmentation in Fig. 9.

3.3. Line spread estimation

The intensity along a line orthogonal to the edge was extracted with the scanner control software on a per pixel basis. This corresponds to an estimate of the edge spread function of the system. The LSF (and by extension PSF) is obtained by computing the differences in intensity. The LSF was normalised such that it sums to 1. Thereafter, the parameters of a Gaussian kernel, estimating the LSF and given by

$$f(x) = \frac{1}{\sigma\sqrt{\pi}} \exp\left(-\frac{(x-\mu)^2}{\sigma^2}\right), \tag{4}$$

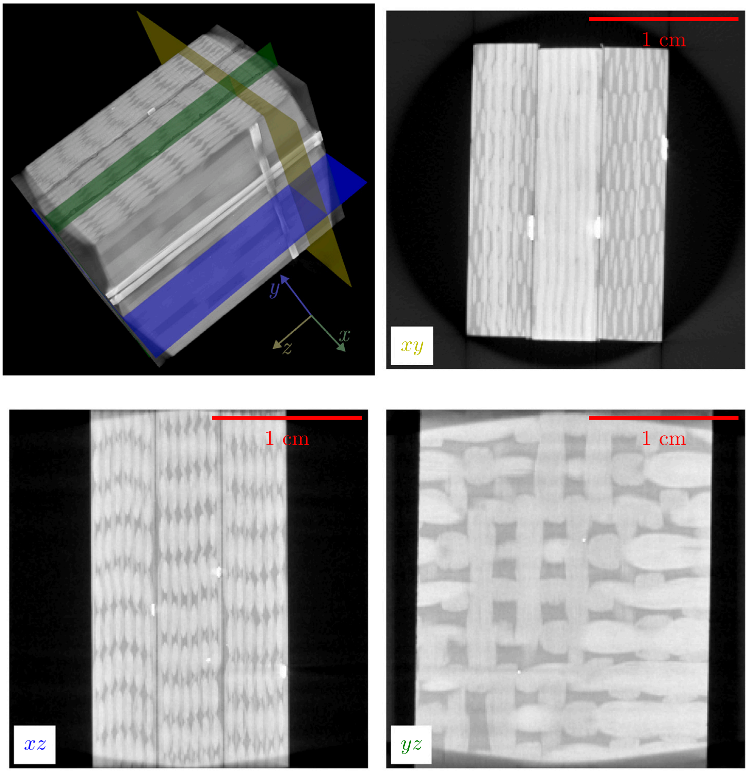


Fig. 8. A 3D view of the XRCT scan of the 3D-woven sample is displayed together with three representative slices (xy-slice no. 100, xz-slice no. 255, and yz-slice no. 320). The very bright areas in xy and xz are due to glass fibre tracer yarns, which are also visible on the surface in the 3D view.

were found through a non-linear least squares fit of the normalised LSF. The normalised measured LSF and the fitted kernel can be seen in Fig. 10. The interesting quantity is the standard deviation σ that determines the width of the kernel. The standard deviation parameter was found to be $\sigma \approx 2.145$ pixels.

This value of the standard deviation pertains the LSF for a binning 1 system. In order to estimate the kernel properties for a binning 4 system a few more steps were carried out. A unit step function was convolved with the computed kernel. This can be viewed as a numerical simulation of an edge spread measurement. The resulting function was thereafter binned as

$$x_{\text{bin}}[j] = \sum_{i=0}^{3} x[4j+i]/4. \tag{5}$$

The task that remained was to find a Gaussian kernel that once convolved with a unit step resulted in the same edge spread function as the binned simulated edge spread function. This was achieved through a second non-linear least squares fit. The standard deviation of this kernel was finally found to be $\sigma_{\text{Bin}} \approx 0.7920$.

3.4. Generated datasets

Three datasets were generated; a training set with 700 synthetic tomographic volumes, a validation set with 150 synthetic tomographic volumes, and a testing set 150 synthetic tomographic volumes. The three datasets, each of a size of 512 by 512 by 512 voxels, were generated with the same set of simulation parameters. The simulation parameters were chosen as to replicate the scan outlined in Section 2.1. Tungsten was used as the electron target material in the X-ray source, and aluminium was used as a filter material. During the simulated scans, the sample rotation direction was decided to be clock-wise or counter clock-wise randomly with equal probability. The entire data generation process took around 35 h, where 10 parallel processes were used for the geometry generation stage in the pipeline.

The relevant parameters for the geometry generation are listed in Table 3. While the parameters were sampled before each geometry was generated, the node perturbations were performed uniformly node-wise with a magnitude decided by parameters in Table 3. For each sample, the yarn type to use for cutting from the other (boolean subtraction) was decided with equal probability. The X-ray tomography simulation parameters are listed in Table 4. The attenuation properties of the material phases were sampled according to Table 5. As a final note, it should be mentioned that the statistics of the parameters pertaining the weave geometry got slightly biased, since for some collections of parameters TexGen crashed, whereby the scan simulation was subsequently restarted with a new set of randomly sampled parameters.

An example of synthetic tomographic reconstruction from the training set is displayed in Fig. 11. The corresponding segmentation is shown in Fig. 12. Histograms of the intensity values of the reconstructions in the three datasets are shown in Fig. 13, together with the histogram of the real scan of the sample. The three histograms look very similar to each other, which indicates that the datasets are large enough. Furthermore, the histograms show qualitative resemblance to the experimental histogram. The peaks for the matrix and yarns appear in similar locations. The matrix peak is, however, relatively larger,

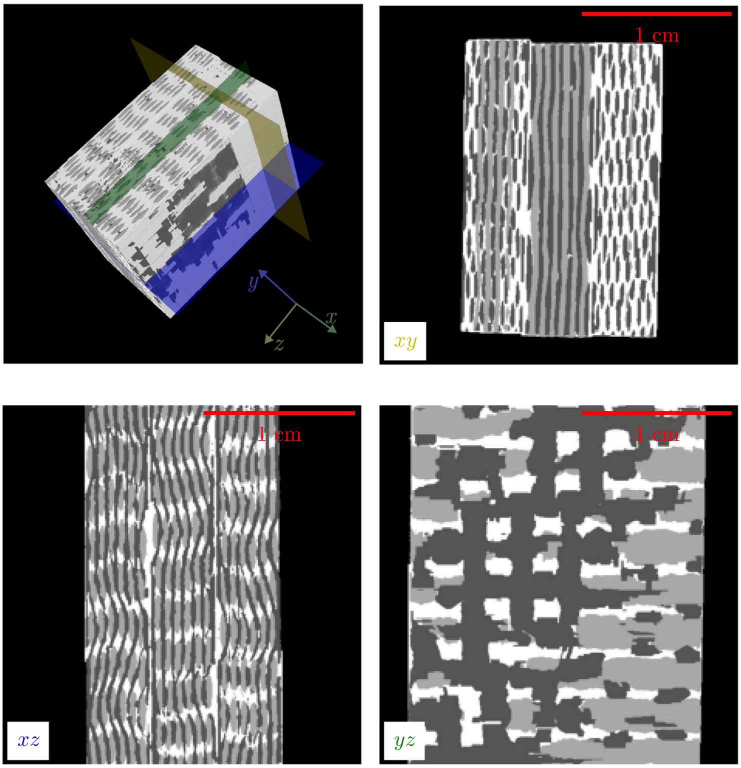


Fig. 9. A 3D view of the manually segmented scan of the 3D-woven sample is displayed together with three representative slices (xy-slice no. 100, xz-slice no. 255, and yz-slice no. 320). Air is represented by black, weft is represented by dark grey, warp is represented by light grey, and pure matrix is represented by white.

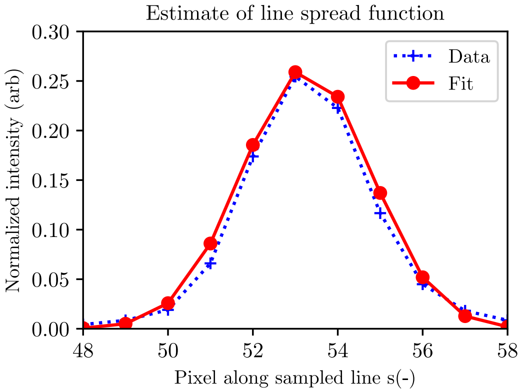


Fig. 10. The differences of the measured edge spread function (Data) is shown with a non-linear least squares fit to a Gaussian kernel (Fit).

which can likely be explained by lower compaction in the simulated geometries and the matrix pockets in the stacking direction, cf. the matrix rich zone in the middle of the xz-slice in Fig. 11 that is not present in Fig. 8.

Table 3
Parameters used in the generation of woven geometries are listed. Parameters were uniformly sampled in these ranges. The cross section magnitudes refers to the range (±) of uniform randomly sampling of node perturbations during textile generation.

Parameter	Lower value	Upper value	Unit
Yarn width to yarn spacing ratio	0.85	0.99	–
Weft to warp ratio	0.35	0.65	–
Number of yarns per layer	3	9	–
Number of yarn layers	5	7	–
Unit cell side length	5.7	22.575	mm
Unit cell thickness	3.8	4.2	mm
Ratio of yarn crossovers	0.05	0.5	–
Tiling	2	3	–
Cross section offset magnitude	0.025	0.075	mm
Cross section scaling magnitude	2.5	7.5	%
Cross section rotation magnitude	2.5	7.5	°

3.5. Trained segmentation model

Based on the statistics on the training dataset described in the previous section the z-score normalisation parameters were determined to be a mean of $\mu = 0.207\,72\,\mathrm{cm}^{-1}$ and a standard deviation of $\sigma = 0.259\,10\,\mathrm{cm}^{-1}$. Similarly the class balancing weights were found to be 0.1242 for air, 1.0038 for weft, 1.0000 for warp, and 0.6890 for matrix. A summary of all used model training hyper parameters can be seen in Table 6. The model was trained on slices from the training set along

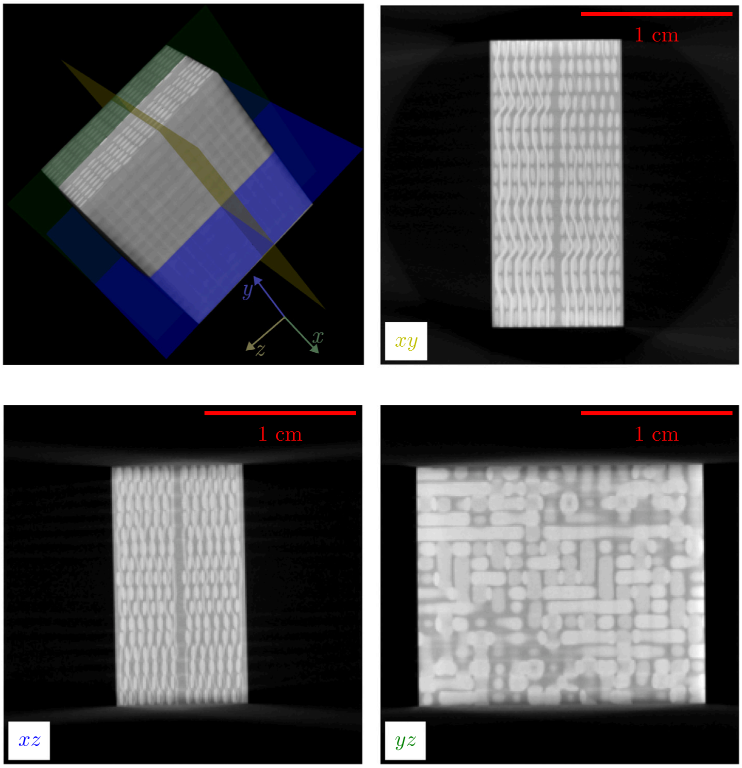


Fig. 11. A 3D view of the simulated XRCT scan of a 3D-woven sample (sample 100 from the training set) is displayed together with three slices (xy-slice no. 236, xz-slice no. 236, and yz-slice no. 236).

the x-axis. That is, the yz-slice in Figs. 11 and 12 would correspond to an input and ground truth, respectively.

The first training stage took 74 h and 20 min, while the second training stage took 7 h and 42 min. The validation loss was calculated online, where the loss for validation batch was computed after each SGD iteration. As the validation set is smaller, the validation set was cycled through 5 times during one training epoch.

A lin-log plot of the loss as a function of iteration number can be seen in Fig. 14. Exponential smoothing is applied to the data for plotting, where the output y at point n is computed from the data at the current point and the smoothed data at the previous point according to

$$y[n+1] = x[n]\alpha + (1-\alpha)y[n]. \tag{6}$$

Here a value $\alpha = 0.0001$ was used as the smoothing factor.

To assess the accuracy of the trained segmentation model, it is used to perform inference on the testing set. As quantitative metrics, we consider the loss, the Jaccard (intersection over union) index, and the pixel-wise accuracy, all computed over the test set (minimum and maximum values refer to individual slices). The Jaccard index is here computed over all material classes for a segmented image A_{ijk} and a ground truth B_{ijk} as

$$J(A,B) = \frac{\sum_{i=1}^{n_x}\sum_{j=1}^{n_y}\sum_{k=1}^{n_c} A_{ijk}B_{ijk}}{\sum_{i=1}^{n_x}\sum_{j=1}^{n_y}\sum_{k=1}^{n_c} A_{ijk}+B_{ijk}-A_{ijk}B_{ijk}}, \tag{7}$$

where A_{ijk} and B_{ijk} are equal to one if the segmented pixel with coordinates i,j is of class k and zero otherwise, and where n_c, n_x, and n_y are the number of classes, the image width, and image height respectively. In addition, the pixel-wise accuracy is defined as the number of correctly classified pixels in an image divided by the total number of pixels as

$$P(A,B) = \frac{1}{n_x n_y}\sum_{i=1}^{n_x}\sum_{j=1}^{n_y}\sum_{k=1}^{n_c} A_{ijk}B_{ijk}. \tag{8}$$

The resulting average, minimum and maximum metrics are summarised in Table 7. As can be seen, all metrics indicate a very high accuracy of the model when used to segment virtually generated geometries.

3.6. Validation on real sample

To validate the trained segmentation model, it is in a second step used to perform a segmentation of the reconstruction of the described physical angle-interlock sample. The segmentation is performed slice-wise along the x-axis, and is concatenated to a volume after each slice segmentation. The segmentation of the entire volume takes around 10 s, and the results are displayed in Fig. 15.

The performance metrics introduced in the previous section are computed for the machine learning based segmentation, using the manual segmentation as ground truth. All metrics are listed in Table

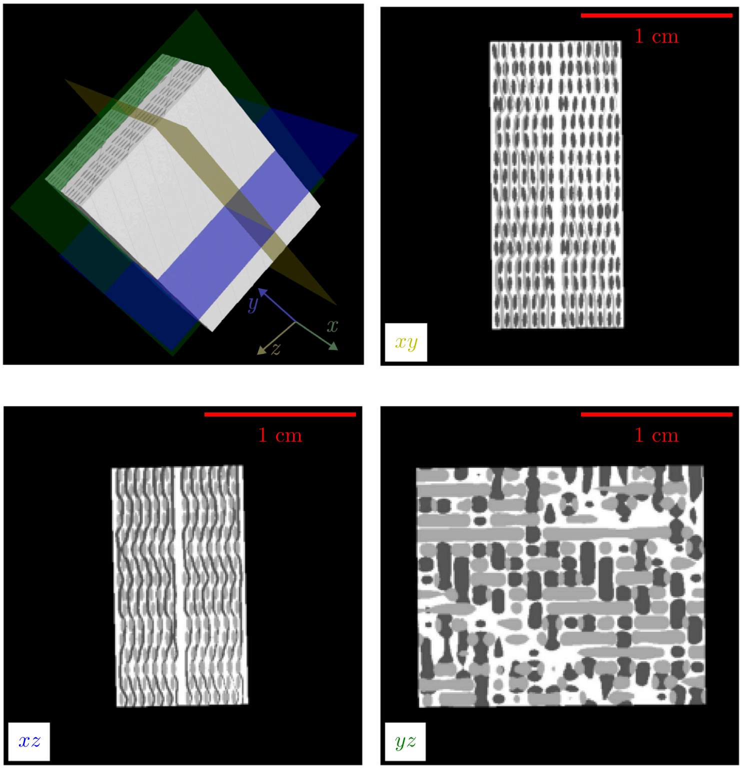


Fig. 12. A 3D view of an automatically performed segmentation of a generated sample (sample 100 from the training set) is displayed together with three slices (xy-slice no. 236, xz-slice no. 236, and yz-slice no. 236).

Table 4
Parameters pertaining the X-ray simulation. "Detector number of rows", "Detector number of columns", and "Binning" were kept constant in order to have uniform image dimensions in the datasets. The remaining parameters were uniformly sampled in the indicated ranges. Note that the detector pixel size is computed to be equivalent to a 0.4 times magnification objective.

Parameter	Lower value	Upper value	Unit
X-ray tube voltage	38	42	kV
X-ray tube power	9.5	0.5	W
X-ray source to sample distance	76	84	mm
Sample to detector distance	142.5	157.5	mm
Sample offset along x-axis	−0.2625	0.2625	mm
Sample offset along y-axis	−0.2625	0.2625	mm
Sample offset along z-axis	−0.2625	0.2625	mm
Sample offset around x-axis	−1.575°	1.575	°
Sample offset around y-axis	−1.575	1.575	°
Sample offset around z-axis	−1.575	1.575	°
Number of projections	2851	3151	–
Exposure time	4.75	5.25	s
Point spread std. deviation (binned)	0.7524	0.8316	–
Scanning angle	342	378	°
Detector pixel size	32.0625	35.4375	μm
Detector number of rows	2048	2048	–
Detector number of columns	2048	2048	–
Binning	4	4	–
Anode angle	11.4	12.6	°
X-ray filter thickness	0.95	1.05	mm

8. It can be noted that the average Jaccard index over the entire volume is above 0.8, and the pixel-wise accuracy is around 0.88. These metrics are comparable with metrics reported in the original model paper [60] and the PyTorch officially reported model benchmarks [62]. Interestingly, the weighted loss metric shows poor performance even though the Jaccard index and pixel-wise accuracy are very high.

Table 5
Parameters relating to the computation of attenuation properties. Parameters were sampled uniformly in these ranges.

Parameter	Lower value	Upper value	Unit
Fibre density	1.691	1.869	$\mathrm{g\,cm^{-3}}$
Fibre diameter	4.94	5.46	μm
Weft fibre count	11,400	12,600	–
Warp fibre count	22,800	25,200	–
Matrix density	1.083	1.197	$\mathrm{g\,cm^{-3}}$
Matrix mixing ratio A : ($B_1 + B_2$)			g : g
A ($C_{25}H_{30}N_2O_4$)	95.0	105.0	g
B_1 ($C_{17}H_{22}N_2$)	32.3475	35.7525	g
B_2 ($C_{21}H_{30}N_2$)	32.3475	35.7525	g
Weft voxels per cross section	190	210	–
Warp voxels per cross section	280	320	–
Voxel area	46.9	46.9	μm^2

3.6.1. *Ability of a model trained on synthetic data to generalise*

As a way to test the robustness of the suggested pipeline, the model is used to segment a sample with a completely different architecture from the one it was trained on. For this purpose a non-woven orthogonal textile by Fureho AB is used. The composite constituents are a RTM6 matrix reinforced with T700GC [63] carbon fibre yarns. The yarns have

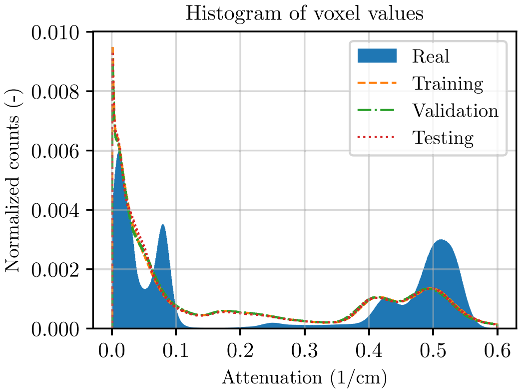


Fig. 13. The intensity histogram of the three datasets are displayed together with the experimental histogram data. Note that values smaller than $1e-6$ are left out, as a bin including such values is an order of magnitude larger than the other bins due to high concentration of near zero values from a cone beam artifact. The rightmost peak corresponds to the yarns, and the second peak from the right corresponds to the matrix.

Table 6
The used model training hyper parameters are listed. Values in parenthesis denote changes for the second training stage.

Parameter	Value
Optimiser	Stochastic gradient descent
Base learning rate	0.001 (0.0001)
Momentum	0.99
Batch size	4
Number of epochs	10 (1)
Weight decay	0.00001
Loss function	Weighted cross-entropy
Loss class weights	
0, Air	0.1242
1, Weft	1.0038
2, Warp	1.0000
3, Matrix	0.6890
Learn rate scheduler	Cosine annealing (Linear decay)
Learn rate scheduler period	2 epochs (N/A)
Training set mean	0.20772
Training set std. deviation	0.25910

Table 7
Evaluation metrics computed over the testing set. A small number is desired for loss, while for the two other metrics a higher number is better.

Metric	Average	Minimum	Maximum
Weighted cross-entropy loss	0.02066	0.4×10^{-5}	0.32373
Jaccard index	0.97808	0.48267	1.00000
Pixel-wise accuracy	0.98840	0.65106	1.00000

Table 8
Evaluation metrics computed on the machine learning segmented real XRCT sample. The manual segmentation is used as the ground truth. A small number is desired for loss, while for the two other metrics a higher number is better.

Metric	Average	Minimum	Maximum
Weighted cross-entropy loss	0.49837	$7.72851 \cdot 10^{-6}$	1.54454
Jaccard index	0.80868	0.49614	1.00000
Pixel-wise accuracy	0.88130	0.65409	1.00000

12,000 fibre count, where the fibres have a diameter of 7 μm, and a density of $1.80\,\mathrm{g\,cm^{-3}}$. The sample consists of 4 roughly 2 by 2 cm pieces (each containing roughly 4 by 4 unit cells) cut out of a cured sheet, stacked on top of each other.

The sample was scanned with identical scanner settings as the main sample with some minor modifications. That is the source to object distance was 90 mm, the object to detector distance was 110 mm, and 1601 projections were taken. The scan was reconstructed with FDK. The effective voxel size in the reconstruction was 60.75 μm. The reconstructed volume, from which two slices are shown in Fig. 16 (left), was segmented along the z-axis, since slices along this axis were most similar to the training data. For comparison, two slices from the machine learning based segmentation of the same sample are shown in Fig. 16 (right).

Based on these results, it is evident that the model qualitatively manages to segment horizontal yarns, but struggles with the through the thickness vertical reinforcing yarns, especially seen in the top layer of the xz-slice. A plausible explanation for this difficulty is that these vertical yarns have no equivalent geometry represented in the training set. It should also be mentioned that the sample described in this section consists of three yarn types (axial yarns, horizontal yarns, and vertical binding yarns), which is not seen in the training data. Bearing these discrepancies in mind, it is still noteworthy how the model manages to segment the highly compressed wedge-shaped yarns (dark grey patches in the 3rd and 4th layers from the top in the xz-segmentation in Fig. 16) even though nothing similar is present in the training data.

4. Discussion

The proposed pipeline was used to generate a synthetic, automatically labelled dataset, which in turn was used to train a segmentation model. It has been demonstrated that this model can perform segmentations that perform in line with models trained on real, manually labelled data. Thereby, the pipeline enables the geometric categorisation of textile composites in their as-manufactured state, without the need of manually labelling slices for the training of a segmentation algorithm. The flexibility and speed of the pipeline also opens up for the possibility of creating more advanced segmentation models that can generalise to segment scans of a variety of weave types performed on a variety of XRCT equipment.

One issue with the segmentation model in its current state is the lack of continuity constraints for yarns. Since the segmentation is performed slice by slice, without any information being propagated between slices, there is no guarantee that the number of yarns is preserved throughout the volume (as is the case in the physical sample). An indication of this can be seen in Fig. 15, where, for example, the bottom of the xz-slice clearly displays weft yarns that become discontinuous.

When comparing the yz-slices in Figs. 8 and 15, it is evident that the segmentation algorithm struggles in the band from the upper right to the lower left corner, which is weft-dominated in the manual segmentation (dark grey). The synthetic woven geometries are, even though randomly perturbed, quite idealised. The chosen geometry generation scheme is not able to generate weaves that display the same level of compaction, and by extension yarn volume fraction, as seen in the real XRCT images. The synthetic geometries also do not show realistic contact between yarns as they are generated to not touch, and where they erroneously touch they are cut out. A future possible improvement to the pipeline is a replacement of the geometric modeller in favour of a more physically based yarn weave model.

The fact that the trained model manages to segment the extra testing sample quite well, except for the vertical yarns, indicates that the fidelity of the X-ray physics modelling is adequate. If the physics model is to be improved one obvious addition is the energy dependent detector response. It would also be of interest to investigate the effect of considering non-homogeneous and/or anisotropic yarns, where the X-ray attenuation depends on the yarn orientation. Such an effect would enable the synthetic data to reproduce the striation patterns observable in real yarns.

It is observed that the class-weighted cross-entropy loss is poor on the real sample validation, even though the other accuracy metrics are resulting in satisfactory numbers. This may be an indication that

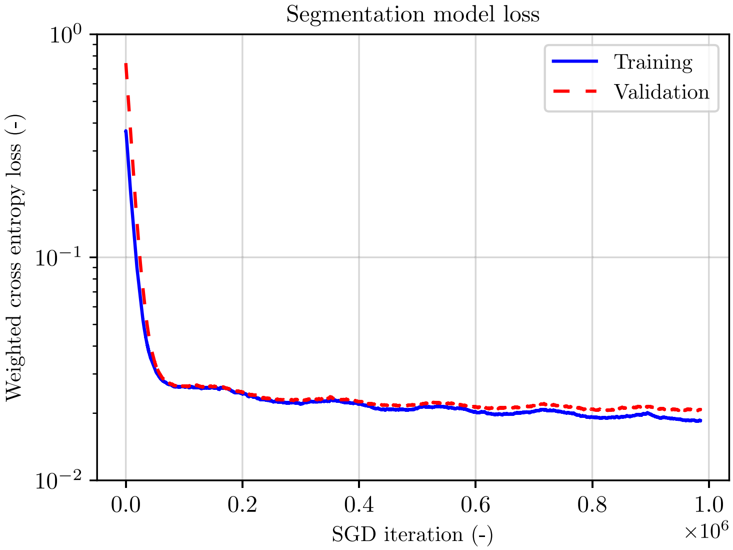


Fig. 14. A lin-log plot of the segmentation algorithm training loss is shown. Exponential smoothing has been applied to the data in the plot.

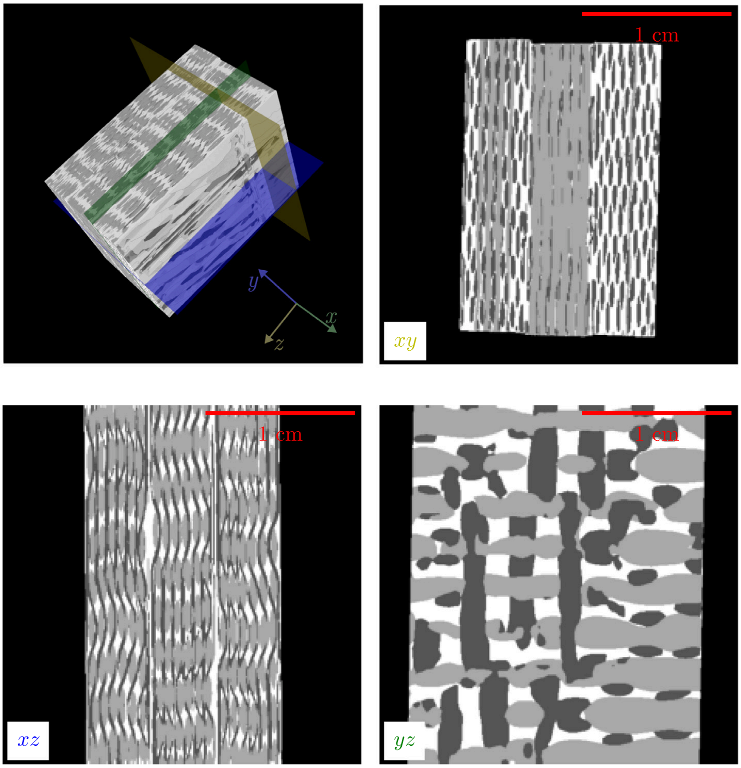


Fig. 15. A 3D view of the machine learning segmented scan of the 3D-woven sample is displayed together with three representative slices (xy-slice no. 100, xz-slice no. 255, and yz-slice no. 320). Air is represented by black, weft is represented by dark grey, warp is represented by light grey, and pure matrix is represented by white. Cf. the original CT images in Fig. 8 and the manual segmentation of the same in Fig. 9.

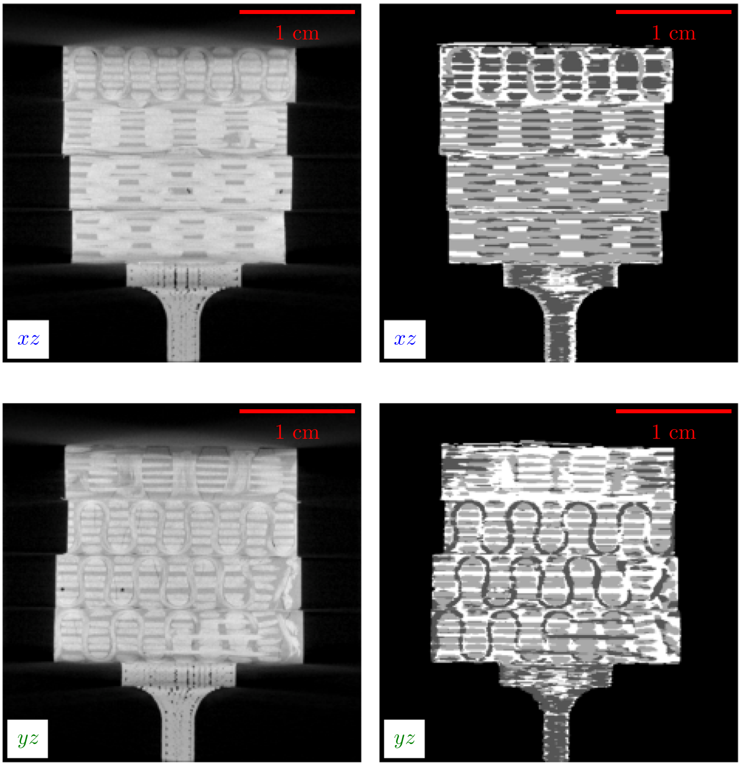


Fig. 16. Two different reconstructed cross sections of the extra testing sample (left) are shown next to their segmentations (right). The sample consists of four stacked plates each roughly 2 by 2 centimetres. The structure visible under the bottom layer is a 3D-printed mount used during XRCT acquisition.

another loss function, such as $1 - J$ where J is the Jaccard index, may improve network training speed and segmentation performance further. An interesting future point of investigation is therefore to compare how choice of loss function affects the ability of a model trained on synthetic data to segment a real XRCT scan. Furthermore, a large scale training hyper parameter optimisation scheme could increase the model viability further.

It is important to circle back to the reason why an XRCT segmentation is being performed in the first place. The end goal is translating the segmentation into a computational mesh, and performing finite element analyses that can predict the mechanical properties of 3D woven CFRPs in their as-manufactured state. The elastic properties can be predicted satisfactory with a voxel mesh. Before any simulation can be performed though, the fibre orientation information need to be mapped to each voxel. This information can for example be extracted from the XRCT data through structure tensor analysis and mapped to individual elements [64]. A very low resolution scan, where fibre structure is not resolved, would however require exploiting other geometric information and introduce new model assumptions.

However, a voxel mesh approach proves unsuitable were local stress concentrations are important to capture, such as when modelling plasticity and damage. For such models a conformal mesh may be required. There are two natural extensions to the pipeline to yield a conformal mesh. One would be to train a segmentation algorithm on geometrical descriptors of a woven structure instead of using the voxelised segmentation. The other is to implement a post processing stage where the semantic segmentation becomes an instance segmentation (individual yarns), and the voxelised geometry is meshed smoothly through an algorithm such as marching cubes [65].

5. Conclusion

In this work a fully open source pipeline for the generation of synthetic, automatically labelled data aimed at training a machine learning model for segmentation of CT-scans of 3D-woven composites is presented. The training data creation requires minimal human intervention. A large labelled dataset, with a thousand reconstructions of a variety of weaves, can be created on a personal computer in the same time it takes to capture just 7–9 comparable real XRCT scans without labelling. Considering that the additional labelling of every single real scan, which if done by hand, would require several person weeks of manual labour, the amount of training data that can be generated in reasonable time is comparatively huge. Thereby, generating model training data this way appears to be a highly efficient and cost-effective alternative.

However, a machine learning-based segmentation model trained entirely on synthetic data is only useful if it manages perform with high accuracy when applied to data from real scans. Therefore, the synthetic data was used to train a prebuilt segmentation model, the `fcn_resnet50` comprising a Fully Convolutional Network segmentation head and a 50 layer ResNet backbone. The results obtained when the trained model is used to segment the output of a XRCT scan of a 3D-woven angle-interlock carbon fibre epoxy composite shows great promise. From these results, it is demonstrated that it is indeed possible to successfully use a machine learning model, entirely trained on synthetic data, to segment low contrast XRCT images of 3D-woven composites to an acceptable accuracy. The segmentation model yields segmented tomographic volumes that largely agree with a segmentation performed manually by a human. More specifically, the machine learning based model displays an average Jaccard index above 0.8, while the pixel-wise accuracy is around 0.88.

To also make a first assessment of how well the trained segmentation model can generalise to unseen geometrical features, the model was also used to segment the output of an XRCT scan of a second type of 3D fabric reinforced carbon fibre epoxy composite. Even though the internal reinforcement architecture of the second sample is quite different from architectures seen during model training, it was demonstrated that the model qualitatively can segment and capture characteristic features also in this case. However, the model had difficulties to capture some features, primarily the existence of additional vertical reinforcement yarns not seen while training.

Overall, it is thereby concluded that it is not only possible, but also effective, to train a machine learning-based segmentation model on 100% synthetic, automatically labelled data. Based on the type of errors observed when segmenting the second type of sample, with explicit features not seen during model training, it is hypothesised that increasing the realism of the internal reinforcement architecture is the main area of possible improvement.

A major implication of this work is the large time savings in XRCT segmentation tasks, which potentially could accelerate material characterisation of woven composites in their as-manufactured state. The proposed pipeline already presents an opportunity to segment poor contrast XRCT reconstructions of 3D-woven CFRPs. A natural next step is to build computational models (e.g. a finite element model) from segmented volumes of different scanned samples, and use these computational models to make predictions of how mechanical properties 3D-woven composites vary with respect to internal variations in the as-manufactured internal reinforcement architecture.

As the segmentation of XRCT volumes are an inherently 3D problem, it is also of interest to investigate segmentation models that exploit volumetric information. For a future study, it is reasonable to assume that even greater segmentation performance can be achieved by upgrading to a 3D-aware model. This could for example be a 3D-convolutional network. Another interesting alternative could be an auto-regressive model, where the prediction of one slice is used as an input for the following slice. This has successfully been used for a residual U-Net, but for a time dependent fluid flow problem [66]. In an auto regressive setting, it can be guaranteed that the number of yarns stays constant, even in low contrast areas.

CRediT authorship contribution statement

Johan Friemann: Writing – original draft, Visualization, Validation, Software, Methodology, Investigation, Formal analysis. **Lars P. Mikkelsen:** Writing – review & editing, Supervision, Resources. **Carolyn Oddy:** Writing – review & editing, Supervision, Resources. **Martin Fagerström:** Writing – review & editing, Supervision, Resources, 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

The authors acknowledge funding from Horizon Europe through the MSCA Doctoral Network RELIANCE: Reliance: REaL-tIme characterisation of ANisotropic Carbon-based tEchnological fibres, films and composites, grant no. 101073040. The X-ray scans were captured at DTU Center for Advanced Structural and Material Testing (CASMaT) located at the 3D imaging Centre (3DIM).

The authors would like to thank Associate Professor Ida Häggström at Chalmers University of Technology, for enlightening discussions and advice on training segmentation models. The authors would also like to extend our thanks to Professor Gunnar Westman at Chalmers University of Technology, for assistance with interpreting the molecular composition of the studied polymers, and computing their atomic mass fractions.

Data availability

Data that is not explicitly referenced as available in the manuscript will be made available on request.

References

[1] Stig F, Hallström S. Assessment of the mechanical properties of a new 3D woven fibre composite material. Compos Sci Technol 2009;69(11–12):1686–92. http://dx.doi.org/10.1016/j.compscitech.2008.04.047.

[2] Guénon VA, Chou TW, Gillespie JW. Toughness properties of a three-dimensional carbon-epoxy composite. J Mater Sci 1989;24(11):4168–75. http://dx.doi.org/10.1007/BF01168991.

[3] Mouritz AP, Baini C, Herszberg I. Mode I interlaminar fracture toughness properties of advanced textile fibreglass composites. Compos Part A Appl Sci Manuf 1999;30(7):859–70. http://dx.doi.org/10.1016/S1359-835X(98)00197-3.

[4] Potluri P, Hogg PJ, Arshad M, Jetavat D, Jamshidi P. Influence of fibre architecture on impact damage tolerance in 3D woven composites. Appl Compos Mater 2012;19(5):799–812. http://dx.doi.org/10.1007/s10443-012-9256-9.

[5] Kazemianfar B, Esmaeeli M, Nami MR. Experimental investigation on response and failure modes of 2D and 3D woven composites under low velocity impact. J Mater Sci 2020;55(3):1069–91. http://dx.doi.org/10.1007/s10853-019-04096-1.

[6] Brown LP, Long AC. Modeling the geometry of textile reinforcements for composites: TexGen. In: Boisse P, editor. Composite reinforcements for optimum performance. Woodhead publishing series in composites science and engineering, 2nd ed.. Woodhead Publishing; 2020, p. 237–65. http://dx.doi.org/10.1016/B978-0-12-819005-0.00008-3.

[7] Lomov SV. Modeling the geometry of textile composite reinforcements: WiseTex. In: Composite reinforcements for optimum performance. 2nd ed.. Elsevier Ltd.; 2020, p. 199–236. http://dx.doi.org/10.1016/B978-0-12-819005-0.00007-1.

[8] Wang Y, Sun X. Digital-element simulation of textile processes. Compos Sci Technol 2001;61(2):311–9. http://dx.doi.org/10.1016/S0266-3538(00)00223-2.

[9] Zhou G, Sun X, Wang Y. Multi-chain digital element analysis in textile mechanics. Compos Sci Technol 2004;64(2):239–44. http://dx.doi.org/10.1016/S0266-3538(03)00258-6.

[10] Mahadik Y, Hallett SR. Finite element modelling of tow geometry in 3D woven fabrics. Compos Part A Appl Sci Manuf 2010;41(9):1192–200. http://dx.doi.org/10.1016/j.compositesa.2010.05.001.

[11] Green SD, Long AC, El Said BS, Hallett SR. Numerical modelling of 3D woven preform deformations. Compos Struct 2014;108(1):747–56. http://dx.doi.org/10.1016/j.compstruct.2013.10.015.

[12] El Said BS, Green S, Hallett SR. Kinematic modelling of 3D woven fabric deformation for structural scale features. Compos Part A Appl Sci Manuf 2014;57:95–107. http://dx.doi.org/10.1016/j.compositesa.2013.11.006.

[13] Auenhammer RM, Mikkelsen LP, Asp LE, Blinzler BJ. Automated X-ray computer tomography segmentation method for finite element analysis of non-crimp fabric reinforced composites. Compos Struct 2021;256(August 2020):113136. http://dx.doi.org/10.1016/j.compstruct.2020.113136.

[14] Sinchuk Y, Kibleur P, Aelterman J, Boone MN, Van Paepegem W. Variational and deep learning segmentation of very-low-contrast X-ray computed tomography images of carbon/epoxy woven composites. Materials 2020;13(4). http://dx.doi.org/10.3390/ma13040936.

[15] Tang C, Zou J, Xiong Y, Liang B, Zhang W. Automatic reconstruction of closely packed fabric composite RVEs using yarn-level micro-CT images processed by convolutional neural networks (CNNs) and based on physical characteristics. Compos Sci Technol 2024;252(April):110616. http://dx.doi.org/10.1016/j.compscitech.2024.110616.

[16] Ewert A, Drach B, Vasylevskyi K, Tsukrov I. Predicting the overall response of an orthogonal 3D woven composite using simulated and tomography-derived geometry. Compos Struct 2020;243(March):112169. http://dx.doi.org/10.1016/j.compstruct.2020.112169.

[17] Lee SU, Chung S, Park RH. A comparative performance study of several global thresholding techniques for segmentation. Comput Vis Graph Image Process 1990;52(2):171–90. http://dx.doi.org/10.1016/0734-189X(90)90053-X.

[18] Roerdink JB, Meijster A. The watershed transform: definitions, algorithms and parallelization strategies. In: Fundamenta informaticae, vol. 41, IOS Press; 2000, p. 187–228. http://dx.doi.org/10.3233/fi-2000-411207.

[19] Naouar N, Vidal-Sallé E, Schneider J, Maire É, Boisse P. Meso-scale FE analyses of textile composite reinforcement deformation based on X-ray computed tomography. Compos Struct 2014;116(1):165–76. http://dx.doi.org/10.1016/j.compstruct.2014.04.026.

[20] Naouar N, Vidal-Sallé E, Schneider J, Maire É, Boisse P. 3D composite reinforcement meso F.E. analyses based on X-ray computed tomography. Compos Struct 2015;132:1094–104. http://dx.doi.org/10.1016/j.compstruct.2015.07.005.

[21] Fourrier G, Rassineux A, Leroy FH, Hirsekorn M, Fagiano C, Baranger E. Automated conformal mesh generation chain for woven composites based on CT-scan images with low contrasts. Compos Struct 2023;308(August 2022):116673. http://dx.doi.org/10.1016/j.compstruct.2023.116673.

[22] Bénézech J, Couégnat G. Variational segmentation of textile composite preforms from X-ray computed tomography. Compos Struct 2019;230(August):111496. http://dx.doi.org/10.1016/j.compstruct.2019.111496.

[23] Pidou-Brion V, Le Guilloux Y. Active yarn meshes for segmentation on X-ray computed tomography of textile composite materials at the mesoscopic scale. Compos Struct 2022;281(December 2021). http://dx.doi.org/10.1016/j.compstruct.2021.115084.

[24] Pannier Y, Coupé P, Garrigues T, Gueguen M, Carré P. Automatic segmentation and fibre orientation estimation from low resolution X-ray computed tomography images of 3D woven composites. Compos Struct 2023;318(April):117087. http://dx.doi.org/10.1016/j.compstruct.2023.117087.

[25] Blusseau S, Wielhorski Y, Haddad Z, Velasco-Forero S. Instance segmentation of 3D woven fabric from tomography images by deep learning and morphological pseudo-labeling. Compos Part B: Eng 2022;247(March). http://dx.doi.org/10.1016/j.compositesb.2022.110333.

[26] Mendoza A, Trullo R, Wielhorski Y. Descriptive modeling of textiles using FE simulations and deep learning. Compos Sci Technol 2021;213(March):108897. http://dx.doi.org/10.1016/j.compscitech.2021.108897.

[27] Zheng K, Wu C, Chen H, Zhang X, Wang Z, Pan Z, et al. Improved 3D image segmentation for X-ray tomographic data of biaxial warp-knitted composites. J Reinf Plast Compos 2023. http://dx.doi.org/10.1177/07316844231220716.

[28] Ali MA, Guan Q, Umer R, Cantwell WJ, Zhang TJ. Deep learning based semantic segmentation of μCT images for creating digital material twins of fibrous reinforcements. Compos Part A: Appl Sci Manuf 2020;139(August):106131. http://dx.doi.org/10.1016/j.compositesa.2020.106131.

[29] Konopczynski TK. Deep learning segmentation algorithms for X-ray CT data [Ph.D. thesis], Heidelberg University; 2021, http://dx.doi.org/10.11588/HEIDOK.00030239.

[30] Ali MA, Guan Q, Umer R, Cantwell WJ, Zhang T. Efficient processing of μCT images using deep learning tools for generating digital material twins of woven fabrics. Compos Sci Technol 2022;217(13):109091. http://dx.doi.org/10.1016/j.compscitech.2021.109091.

[31] Tsamos A, Evsevleev S, Fioresi R, Faglioni F, Bruno G. Synthetic data generation for automatic segmentation of X-ray computed tomography reconstructions of complex microstructures. J Imaging 2023;9(2). http://dx.doi.org/10.3390/jimaging9020022.

[32] Fok WYR, Fieselmann A, Huemmer C, Biniazan R, Beister M, Geiger B, et al. Adversarial robustness improvement for X-ray bone segmentation using synthetic data created from computed tomography scans. Sci Rep 2024;14(1):25813. http://dx.doi.org/10.1038/s41598-024-73363-2.

[33] Friemann J. Textomos [software]. 2025, https://github.com/Johan-Friemann/textomos.

[34] Friemann J. X-ray CT scan of layer to layer angle interlock composite sample with hand segmentation [dataset]. Zenodo; 2025, http://dx.doi.org/10.5281/zenodo.14891845.

[35] Friemann J. X-ray CT scan of orthogonal noobed composite sample. Zenodo; 2025, http://dx.doi.org/10.5281/zenodo.15389427.

[36] Hexcell. HexTow® IM7 carbon fiber. Hexcell; 2023, p. 1–3, Product data sheet, https://www.hexcel.com/user_area/content_media/raw/IM7_HexTow_DataSheet.pdf.

[37] Hexcell. HexFlow® RTM6/RTM6-2 180 °C epoxy system for resin transfer molding and infusion technologies. Hexcell; 2023, p. 1–3, Product data sheet, https://www.hexcel.com/user_area/content_media/raw/RTM6_RTM62_HexFlow_DataSheet.pdf.

[38] Appleby-Thomas GJ, Hazell PJ, Stennett C. The variation in lateral and longitudinal stress gauge response within an RTM 6 epoxy resin under one-dimensional shock loading. J Mater Sci 2009;44(22):6187–98. http://dx.doi.org/10.1007/s10853-009-3859-z.

[39] Parsons AJ, Gonciaruk A, Zeng X, Thomann FS, Schubel P, Lorrillard J, et al. Controlling mass loss from RTM6 epoxy resin under simulated vacuum infusion conditions. Polym Test 2022;107(November 2021):107473. http://dx.doi.org/10.1016/j.polymertesting.2022.107473.

[40] Westman G. Private communication professor at division of chemistry and biochemistry, department of chemistry and chemical engineering. Chalmers University of Technology; 2024.

[41] Oddy C, yi Song M, Stewart C, El Said B, Ekh M, Hallett SR, et al. Cyclic behaviour of 3D-woven composites in tension: Experimental testing and macroscale modelling. Compos Part A: Appl Sci Manuf 2024;187:108354. http://dx.doi.org/10.1016/j.compositesa.2024.108354.

[42] Feldkamp LA, Davis LC, Kress JW. Practical cone-beam algorithm. J Opt Soc Amer A 1984;1(6):612–9. http://dx.doi.org/10.1364/JOSAA.1.000612.

[43] Chen Z, Calhoun VD, Chang S. Compensating the intensity fall-off effect in cone-beam tomography by an empirical weight formula. Appl Opt 2008;47(32):6033–9. http://dx.doi.org/10.1364/AO.47.006033.

[44] van Aarle W, Palenstijn WJ, Cant J, Janssens E, Bleichrodt F, Dabravolski A, et al. Fast and flexible X-ray tomography using the ASTRA toolbox. Opt Express 2016;24(22):25129–47. http://dx.doi.org/10.1364/OE.24.025129.

[45] Fedorov A, Beichel R, Kalpathy-Cramer J, Finet J, Fillion-Robin J-C, Pujol S, et al. 3D slicer as an image computing platform for the quantitative imaging network. Magn Reson Imaging 2012;30(9):1323–41. http://dx.doi.org/10.1016/j.mri.2012.05.001.

[46] Horvath D, Erdos G, Istenes Z, Horvath T, Foldi S. Object detection using Sim2Real domain randomization for robotic applications. IEEE Trans Robot 2023;39(2):1225–43. http://dx.doi.org/10.1109/TRO.2022.3207619.

[47] Cignoni P, Callieri M, Corsini M, Dellepiane M, Ganovelli F, Ranzuglia G. MeshLab: An open-source mesh processing tool. In: 6th Eurographics Italian chapter conference 2008 - proceedings. 2008, p. 129–36.

[48] Zhou Q, Grinspun E, Zorin D, Jacobson A. Mesh arrangements for solid geometry. ACM Trans Graph 2016;35(4). http://dx.doi.org/10.1145/2897824.2925901.

[49] Vidal FP, Garnier M, Freud N, Létang JM, John NW. Simulation of X-ray attenuation on the GPU. In: Theory and practice of computer graphics 2009, TPCG 2009 - Eurographics UK chapter proceedings. 2009, p. 25–32. http://dx.doi.org/10.2312/LocalChapterEvents/TPCG/TPCG09/025-032.

[50] Vidal FP, Villard PF. Development and validation of real-time simulation of X-ray imaging with respiratory motion. Comput Med Imaging Graph 2016;49:1–15. http://dx.doi.org/10.1016/j.compmedimag.2015.12.002.

[51] Freud N, Duvauchelle P, Létang JM, Babot D. Fast and robust ray casting algorithms for virtual X-ray imaging. Nucl Instrum. Methods Phys Res Sect B: Beam Interact. Mater Atoms 2006;248(1):175–80. http://dx.doi.org/10.1016/j.nimb.2006.03.009.

[52] Pointon JL, Wen T, Tugwell-Allsup J, Sújar A, Létang JM, Vidal FP. Simulation of X-ray projections on GPU: Benchmarking gVirtualXray with clinically realistic phantoms. Comput Methods Programs Biomed 2023;234:107500. http://dx.doi.org/10.1016/j.cmpb.2023.107500.

[53] Schoonjans T, Brunetti A, Golosio B, Sanchez Del Rio M, Solé VA, Ferrero C, et al. The xraylib library for X-ray-matter interactions. Recent developments. Spectrochim Acta - Part B At Spectrosc 2011;66(11–12):776–84. http://dx.doi.org/10.1016/j.sab.2011.09.011.

[54] Bushberg JT, Seiberg AJ, Leidholdt EM, Boone JM. The essential physics of medical imaging. Lippincott Williams & Wilkins; 2020.

[55] Poludniowski G, Omar A, Bujila R, Andreo P. Technical note: Spekpy v2.0—a software toolkit for modeling x-ray tube spectra. Med Phys 2021;48(7):3630–7. http://dx.doi.org/10.1002/mp.14945.

[56] Schwarz M, Seidel HP. Fast parallel surface and solid voxelization on GPUs. ACM Trans Graph 2010;29(6):1–10. http://dx.doi.org/10.1145/1882261.1866201.

[57] Pineda J. A parallel algorithm for polygon rasterization. In: proceedings of the 15th annual conference on computer graphics and interactive techniques, SIGGRAPH 1988. 22, (4):1988, p. 17–20. http://dx.doi.org/10.1145/54852.378457.

[58] Okuta R, Unno Y, Nishino D, Hido S, Loomis C. CuPy: A NumPy-compatible library for NVIDIA GPU calculations. In: Proceedings of workshop on machine learning systems (learningSys) in the thirty-first annual conference on neural information processing systems. 2017, URL http://learningsys.org/nips17/assets/papers/paper_16.pdf.

[59] Paszke A, Gross S, Massa F, Lerer A, Bradbury J, Chanan G, et al. PyTorch: An imperative style, high-performance deep learning library. In: Advances in neural information processing systems 32. Curran Associates, Inc.; 2019, p. 8024–35, URL http://papers.neurips.cc/paper/9015-pytorch-an-imperative-style-high-performance-deep-learning-library.pdf.

[60] Shelhamer E, Long J, Darrell T. Fully convolutional networks for semantic segmentation. 2016, arXiv:1605.06211.

[61] He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. In: 2016 IEEE conference on computer vision and pattern recognition. 2016, p. 770–8. http://dx.doi.org/10.1109/CVPR.2016.90.

[62] Team P. FCN: Fully-convolutional network model with ResNet-50 and ResNet-101 backbones. 2025, https://pytorch.org/hub/pytorch_vision_fcn_resnet101/. [Accessed 03 February 2025].

[63] Toray. (T700G)® standard modulus carbon fiber. 2018, Product data sheet, https://www.toraycma.com/wp-content/uploads/T700G-Technical-Data-Sheet-1.pdf.

[64] Auenhammer RM, Jeppesen N, Mikkelsen LP, Dahl VA, Asp LE. X-ray computed tomography data structure tensor orientation mapping for finite element models — Stxae[formula presented]. Softw Impacts 2022;11:100216. http://dx.doi.org/10.1016/j.simpa.2021.100216.

[65] Lorensen WE, Cline HE. Marching cubes: a high resolution 3D surface construction algorithm. Comput Graph (ACM) 1987;21(4):163–9. http://dx.doi.org/10.1145/37402.37422.

[66] Jiang Z, Tahmasebi P, Mao Z. Deep residual U-net convolution neural networks with autoregressive strategy for fluid flow predictions in large-scale geosystems. Adv Water Resour 2021;150(September 2020):103878. http://dx.doi.org/10.1016/j.advwatres.2021.103878.

PAPER B

From X-ray CT to finite element models: A fully automated pipeline for mesoscale modelling of as-manufactured textile composites

Johan Friemann, Carolyn Oddy, Lars P. Mikkelsen, Martin Fagerström

Published in Composites Science and Technology,
Vol. 278, 111561, May. 2026.

https://doi.org/10.1016/j.compscitech.2026.111561

Composites Science and Technology 278 (2026) 111561

Contents lists available at ScienceDirect

Composites Science and Technology

journal homepage: www.elsevier.com/locate/compscitech

From X-ray CT to finite element models: A fully automated pipeline for mesoscale modelling of as-manufactured textile composites

Johan Friemann [a],*, Carolyn Oddy [a,c], Lars P. Mikkelsen [b], Martin Fagerström [a]

[a] Department of Industrial and Materials Science, Chalmers University of Technology, SE-412 96 Gothenburg, Sweden
[b] Composites Manufacturing and Testing, Department of Wind and Energy Systems, Technical University of Denmark, DK-4000 Roskilde, Denmark
[c] GKN Aerospace Sweden, SE-461 81 Trollhättan, Sweden

ARTICLE INFO

Keywords:
X-ray computed tomography
Segmentation
Machine learning
Finite element modelling
3D-textile reinforced composites

ABSTRACT

Structural parts incorporating 3D-textile reinforced composites show great promise in high performance lightweight applications. For widespread industrial use, accurate predictions of mechanical properties are required. Experimental testing campaigns to generate these properties can be prohibitively expensive. To overcome this, meso scale models of the yarn architecture can be derived from X-ray Computed Tomography (XRCT), and computational homogenisation can be performed in the material's as-manufactured configuration. In this work a fully automated pipeline for the prediction of the full 3D elastic properties of 3D-reinforced textile composites from XRCT scans is presented. The proposed methodology enables the study of variations in the as-manufactured material properties from a single large field of view XRCT scan encompassing multiple unit cells, such that finite element homogenisation can yield statistical information. The pipeline includes a machine learning based segmentation model, finite element meshing and boundary condition assignment, and a material mapping procedure. For segmentation, a for textile reinforced composites completely novel 3D U-net architecture can be utilised, owing to the use of a fully synthetic automatically labelled training data set. An application of the pipeline on a 3D-reinforced material sample results in accurately predicted homogenised elastic stiffnesses, with a deviation from experiments of less than 6.5%.

1. Introduction

Composite components incorporating 3D-textiles are gaining attention partly due to their excellent impact and delamination resistance [1]. This, among other factors is important for safety in aerospace where bird strikes, fan blade-offs, and high dynamic loads are ever present risks. Therefore, textile composites with 3D-reinforcements are becoming more widespread in the aerospace industry. The commercially available LEAP engine, for example, utilises 3D-woven carbon fibre reinforced polymers (CFRP) in its fan blades and casing [2]. In order to increase the adoption of 3D-woven composites in the aerospace industry, accurate determination of elastic properties is paramount. For certification purposes it is further motivated to be able to give statistical measures on structural as-manufactured parts. A clear understanding of how material variations influence stiffness and strength, and how these properties are linked to the micro and meso scale material morphology, is essential for component manufacturers to reduce their safety factors. Reduced safety factors will lead to reduced use of material, making components more lightweight while saving raw material and energy.

To elucidate the effect of material morphology variations on mechanical properties, two approaches can be taken: experimental testing or numerical modelling. With physical testing, an excessive amount of samples would have to be manufactured and tested to obtain the necessary statistics. Furthermore, given that the material heterogeneity is typically in the order of tens of millimetres, standard test method for CFRP materials cannot be used. As a consequence, standard test methods have been adjusted such that the gauge regions accommodate a sufficiently large representative volume, cf. e.g. [3–6]. Still, these factors make numerical modelling a more feasible path forward.

In terms of modelling, the material could realistically be considered on two scales: the meso (yarn) scale or the macro (homogenised) scale. Aiming to model large structural parts with high fidelity meso scale models, where reinforced yarns are explicitly represented, are prohibitively computationally expensive. Therefore, a macro scale homogenised model becomes necessary for evaluating structural designs, especially in the prototyping stage of product development. For such models to be reliable, the macro scale properties and their spatial variation have to be used as input.

---

* Corresponding author.
E-mail address: johan.friemann@chalmers.se (J. Friemann).

https://doi.org/10.1016/j.compscitech.2026.111561
Received 8 December 2025; Received in revised form 12 January 2026; Accepted 7 February 2026

Numerically, such equivalent properties are typically derived through finite element (FE) simulations of representative volume elements (RVE), which capture the characteristic periodic structure of the woven material's meso scale. For the purpose of performing RVE simulations, geometry that captures the material morphology accurately needs to be derived. Geometry can be derived through purely geometric means using tools such as TexGen [7] or WiseTex [8]. This method can, however, result in far too idealised geometries that do not capture the as-manufactured properties of the studied material. As a consequence, there has been extensive work on modelling the weave process itself to generate more realistic RVE's. This started with the Digital Element Method (DEM) through Wang et al. [9], refined by Zhou et al. [10] and has been further improved by several different research groups [11–14]. One drawback of this second approach is the extensive CPU time required per unit cell (UC), which makes larger testing campaigns challenging. An alternative high performance method utilising position based dynamics has recently been presented for yarn deformation modelling by Gao et al. [15]. For all these cases, it is however unclear how variability in the meso structure can be accounted for, even more so how it can be linked to variability in part geometry or manufacturing.

That said, in recent years it has become popular to derive computational geometry from x-ray computed tomography (XRCT) scans. A clear benefit is that computational homogenisation of textile composites then can be performed in the textile's as-manufactured state, yielding realistic estimates of the elastic material properties. In fact, components could be scanned at various locations, such that realistic material variability could be acquired and accounted for. Ewert et al. [16] have shown that such geometries can outperform comparable DEM derived geometries.

The main challenge of deriving meso scale RVE geometry from XRCT scans is the segmentation problem. Namely, categorising each voxel in the volume as its corresponding phase (pure matrix, weft yarn, warp yarn, or voids). Due to the significant material heterogeneity, and thereby to capture the entire UC, scans need to include a fairly large field of view. The demands on field of view increase even further if material variability statistics are to be extracted. For efficiency, scanned samples need to include several UCs. Having the desired field of view will inevitably result in poor contrast, making segmentation through classical methods, such as thresholding or watershed, difficult or even impossible [17,18]. Furthermore, assigning local yarn fibre volume fractions and local fibre orientations poses additional challenges.

As alternatives to thresholding or watershed, other researchers have utilised commercial fibre tracking software [19], structure tensor analysis [20–22], or variational methods [23–25] for segmentation. For particularly large field of view scans, which is necessary for cm scale UCs and/or multiple UC scans, these methods might fail. The reason is that they all rely on structural information that may not be resolved in a low contrast scan, or require very long computational times which could make industrial application difficult.

Machine learning based segmentation methods have become increasingly popular for woven composites, and are seeing widespread use [26–28]. A drawback of machine learning based methods is the need for large labelled data sets. There has been work to get around this limitation by using synthetic data. This has for example been done by creating "pseudo CT" by training a surrogate model [29], using "pseudo labels" derived from classical image processing methods [30], or fortifying the data set with synthetic data based on statistics from a manually labelled data set [31]. A fully synthetic data has been used for segmentation of a six-phase Al–Si alloy composite reinforced with ceramic fibres and particles by Tsamos et al. [32]. Furthermore, the current authors recently proposed a fully automated pipeline for training a machine learning based segmentation model entirely using synthetic data [33]. The initial application of this model to segment an XRCT scan of a layer-to-layer angle interlock reinforced composite material showed promise in terms of obtaining acceptable voxel-wise accuracy. A drawback was, however, that continuity of reinforcement yarns could not be guaranteed, something that is being addressed in the current work.

With an accurate segmentation at hand, as-manufactured properties can be derived through computational homogenisation. In this work a fully automated and open source[1] [34] pipeline to generate the necessary finite element models directly from large field of view CT scans of 3D-woven composites is presented. The pipeline extends previous work in utilising fully synthetic training data for machine learning based segmentation of 3D-reinforced textile composites. In addition to segmentation, it now also handles model discretisation, mapping of local fibre orientation and volume fraction, FE input file generation, and elastic property homogenisation. The only manual labour required is the identification of individual UCs from a large field of view segmentation. To promote 3D continuity of the segmented volume, the segmentation problem is handled with a novel 3D-convolutional neural network. Extending the work by Jing et al. [35], who performed meso scale elastic homogenisation of orthogonal weaves using an XRCT guided but manually defined geometry, the current work directly utilises the segmented volume information as basis for the numerical models. Furthermore, the current pipeline is free from any restrictions on (partially) manual data labelling, as done e.g. by Ahmed et al. [36]. Thereby, the pipeline enables a fully automated and efficient generation of multiple UC models from a single scan.

Leveraging the opportunity to automatically generate multiple UC models for macroscopic property extraction, we demonstrate the pipeline on a layer-to-layer angle interlock composite. To study variability of properties and how these are linked to variations in the reinforcement architecture, six unique RVEs are derived from a single large field of view XRCT scan. We demonstrate that macroscopic elastic properties can be predicted in good agreement with experimental observations. As an additional finding, we observe that the variability in macroscopic properties is very small, even when noticeable variations in the reinforcement architecture are present.

2. Material sample and its XRCT scan

The studied material is a 3D-woven layer-to-layer angle interlock CFRP manufactured through resin transfer moulding (RTM). The yarns are made out of IM7 fibres [37] with a diameter of 5.2 μm. The warp and weft yarns are illustrated in blue and red respectively in Fig. 1(a). Warp yarns run in the main weaving direction of the machine and are held in tension during the manufacturing process. The weft yarns on the other hand run transversely to the warp yarns and are inserted during the weaving process. The fibre count is 12k and 24k for the weft and warp yarns, respectively. The matrix consists of RTM6 epoxy [38]. It should be noted that the weave has a quite large UC when compared with the scale of the fibre diameter. Further, it should be noted that the UC of this material is shorter in the warp direction compared to the other two directions. As shown in Fig. 1(a), the UC is 6.5 mm wide in the warp (1) direction, 21.5 mm long in the weft (2) direction, and 4 mm thick in the out-of-plane (3) direction. Fig. 1(a) also denotes and visualises the orientation conventions used in this paper. Further, it should be mentioned that the material contains some glass fibre tracer yarns that were used as aid during manufacturing. Since glass has a much higher x-ray attenuation than carbon fibres and epoxy, this can cause artifacts during XRCT scanning.

For the scan, a sample was created from 3 pieces cut out of a cured panel and stacked on top of each other using double-sided tape, see Fig. 1(b). Each piece contains roughly 2 by 3 UCs. The stacking is performed to obtain a relatively low aspect ratio in the 13-plane of the sample,

---

[1] All data not explicitly referenced, that has been used in this manuscript can be generated with the provided code. If the original data (roughly 700 GB) is needed it will be made available on request.

thereby yielding a relatively uniform x-ray attenuation when scanning. As a benefit, this allows for a large field of view XRCT scan which encompasses as many UCs as possible in a smaller scanning volume.

2.1. XRCT scan and sample reconstruction

The sample was scanned with a *Zeiss XRadia 410 Versa* lab-based XRCT scanner. It uses a *Hamamatsu Photonics L8121-03* micro focus x-ray source with a tungsten target. There is a documented 200 µm beryllium filter inside the source, but initial investigations suggests that there may be an additional undocumented 1 mm aluminium filter inside the source [33]. No additional filter is used. The detector in the imaging system is an *ANDOR iKon-L DW936N-BV-556* CCD detector, with 2 048 by 2 048 pixels with dimensions 13.5 µm by 13.5 µm. One special feature of the Zeiss XRadia 410 Versa is the presence of an optical magnification system between the x-ray scintillation panel and the detector. Here the 0.4 times magnification large field of view objective was used. The source to sample distance was set to 80 mm, while the sample to detector distance was 150 mm. An x-ray tube voltage of 40 kV and power of 10 W was used. The sample was rotated 360° where 3 001 projections, each with an exposure time of 5 s, were taken at equal angular intervals. Binning 4 was performed, which implies that each pixel in the output projections consisted of the average of 4 by 4 pixels. The geometric and objective magnification combined with the binning results in an effective voxel size of 46.96 µm.

The Feldkamp–Davis–Kress (FDK) technique [39], which is a generalisation of filtered back projection for cone beam scans, was used for tomographic reconstruction. The GPU implementation of FDK inside the ASTRA toolbox [40] was used to perform the reconstruction. The reconstructed scan and a manually annotated segmentation of the scan are hosted online [41].

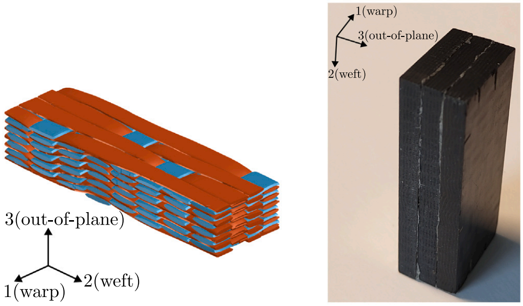
A selection of cross-sectional images of the reconstruction are shown in Fig. 4 (left) below. The combination of the large field of view XRCT scan, together with the lack of contrast between the impregnated yarns and matrix pockets, makes automated segmentation (i.e. the assignment of a material phase to each voxel) challenging using traditional methods. For this reason, a machine learning based segmentation algorithm provides a promising path forward.

3. Machine learning based segmentation of textile reinforced composites

The aim of the study is to develop a fully automated pipeline to predict the full elastic stiffness tensor of textile reinforced composites directly from CT images. As a first stage, segmenting the reconstruction is necessary. This process and its application to the sample introduced in Section 2 is discussed in more detail in the following subsections. It is further highlighted here that all novel contributions to this work are open source [34]. The training of the segmentation model has been carried out on a laptop with an Intel i9-13900H CPU, 32 GB of RAM, a 1 TB NVMe M.2 Performance SSD, and an Nvidia RTX 4080 Laptop GPU (12 GB VRAM).

3.1. Development and training of the machine learning model

Segmentation in this context entails assigning each voxel in the reconstructed XRCT scan to one of four phases: weft yarns, warp yarns, pure matrix, or air. For CFRP composites in particular, the similarity in x-ray attenuation between pure carbon fibres and hydro-carbon epoxy yields volumes with very poor contrast. As a consequence, classical segmentation strategies such as thresholding and watershed are often insufficient. Therefore, machine learning based segmentation models are gaining in popularity, as they can learn to effectively distinguish phase boundaries while ignoring noise and artifacts. The use of machine learning to carry out segmentation requires the generation of labelled 3D training data, the construction of an appropriate neural network architecture and a strategy for the training process.



(a) (b)

Fig. 1. The studied layer-to-layer angle interlock material is shown with the material coordinate system $(1, 2, 3)$. A schematic of the UC is shown in (a) where weft yarns are shown in red, and warp yarns are shown in blue. Original graphics provided by courtesy of Oddy [6]. The specimen scanned in this study is displayed in (b).

When it comes to the generation of labelled 3D training data, one major drawback and criticism of machine learning base approaches is the need for large data sets. This makes the use of traditional hand labelling challenging and exorbitantly time consuming. A possible alternative, and the one considered here, is the use of synthetic training data in which the constituents are automatically labelled. More specifically, in this work a method for the generation of synthetic, automatically labelled training data, recently proposed by the authors [33], is used. Quantitative similarity to real data was demonstrated through comparisons of intensity histograms. Further, the ability of a neural network trained on synthetic data to segment real tomogram slices was demonstrated with 88% pixel-wise accuracy. In contrast to previous work in the field, cf. e.g. Tsamos et al. [32], the synthetic data includes physically motivated attenuation features, noise, and artifacts. Using this method, a data set with 700 volumes used for training, 150 volumes for validation and 150 volumes for testing was generated. The exact same generation parameters used in the previous work were also used here.

Once the training set has been generated, the architecture of the neural network must be considered. In the context of segmenting scans of textile reinforced composites, 2D U-net architectures have been widely used with great success, cf. e.g. Mendoza et al. [29], Zheng et al. [31], Sinchuk et al. [42], Ying et al. [28], and Cao et al. [43]. However, 2D U-net architectures lack the ability to ensure continuity from one volume slice to the next. With the considered aim of generating volumetric model representation of the scanned samples a 3D-convolutional network based on the U-net architecture was considered here. This choice was further supported by the promising results by Tsamos et al. [32] in segmenting an advanced multiphase composite using synthetic training data and a U-net with 3D convolution kernels.

The considered neural network segmentation model is implemented in the PyTorch [44] framework. The implemented architecture is displayed in Fig. 2. The network is constructed from 5 main building blocks. The *Double convolution* block consists of two chained groups consisting of a 3D convolution layer, a batch normalisation layer, and a ReLU activation function. The convolution layers use a kernel size of 3 and a padding of 1. The *Max pool* block consists of a 3D max pool layer that utilises a kernel size of 2 and a stride of 2. The *Up convolution* block refers to a 3D transposed convolution with a kernel size of 2 and stride of 2. The *Concatenation* block concatenates the two inputs channel-wise. The final convolution block is a 3D convolution layer with kernel size

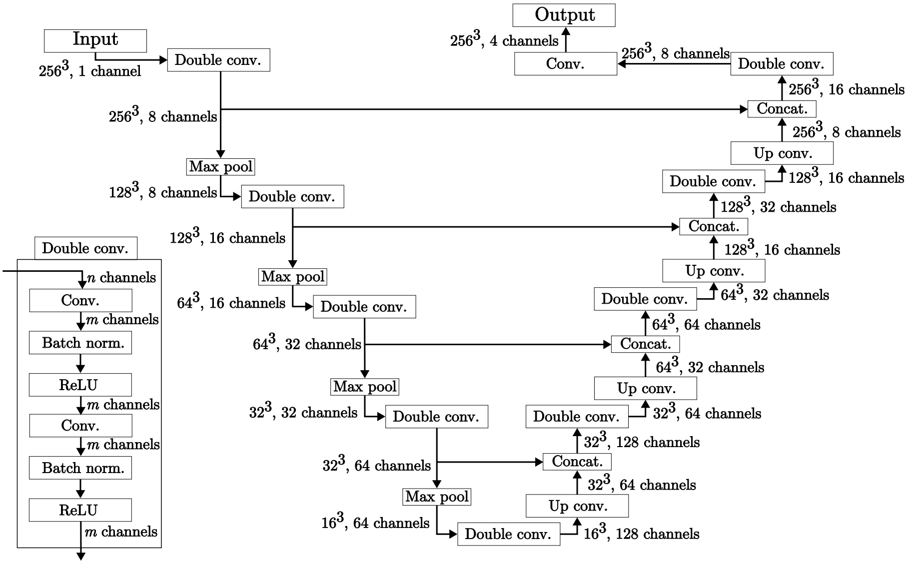


Fig. 2. The used 3D convolution U-net architecture is displayed. The input is a one channel (greyscale) XRCT image and the output with four channels represents the segmentation class probabilities. The structure of the *Double convolution* block is also shown.

1. Its purpose is to yield the desired number of output channels. All parameters not explicitly referred to use the PyTorch default values.

The training process itself requires further pre-processing of the training and validation data sets (splitting and normalisation), the definition of an appropriate loss function, and a selection of a set of important hyperparameters. The synthetic tomographic training volumes were 512 by 512 by 512 voxels in size. The volumes were divided into overlapping patches of 256 by 256 by 256 voxels, with a stride of 128. This means that each tomographic volume results in 27 data points. The input grey scale data is subject to z-score normalisation. This entail subtracting the mean and dividing by the standard deviation, where the statistics has been computed over the entire training set.

A loss function based on the Jaccard index (intersection over union) was used during training. Let A_{ijkl} be a four dimensional array representing a segmented image, which is equal to one if a voxel with coordinates i, j, k is of class l and zero otherwise. The loss for a batch of predicted segmentations A^m_{ijkl} and their corresponding arrays of the ground truths B^m_{ijkl} is defined as

$$J(A,B) = \frac{1}{n_\mathrm{b} n_\mathrm{c}} \sum_{m=1}^{n_\mathrm{b}} \sum_{l=1}^{n_\mathrm{c}} \left(1 - \frac{\sum_{i=1}^{n_x} \sum_{j=1}^{n_y} \sum_{k=1}^{n_z} A^m_{ijkl} B^m_{ijkl} + \epsilon}{\sum_{i=1}^{n_x} \sum_{j=1}^{n_y} \sum_{k=1}^{n_z} (A^m_{ijkl} + B^m_{ijkl} - A^m_{ijkl} B^m_{ijkl}) + \epsilon} \right), \tag{1}$$

where n_c, n_b, n_x, n_y, and n_z are the number of classes, the batch size, the image width, image height, and image depth respectively. The parameter ϵ is here set to 1×10^{-6} to prevent division by zero for data where one or several of the classes are not present.

When training, A is obtained by passing the network output through a channel wise softmax function. This means that in practice the entries A_{ijkl} are not strictly 1 or 0. A training batch size of 2 is used for gradient update computations. The model is trained with the ADAM optimizer for 10 epochs. The validation loss is evaluated online, which means that the loss is computed from a batch of the evaluation set after each gradient update. The validation set is roughly 5 times smaller than the training set. This means that the validation set is repeated about 5 times during one training epoch. The model was trained with a starting learning rate of 0.001. The learning rate is reduced linearly after each gradient update until it reaches 0 at the end of epoch 5. Thereafter, the learning rate is re-set to 0.001, whereafter it again is reduced linearly until it reaches 0 at the end of epoch 10. This "warm" restarting-type of learning rate scheduling has shown potential to improve convergence [45,46]. All other optimizer parameters are kept at the PyTorch default settings.

The model training and validation loss curves are displayed in Fig. 3. The loss does not show any signs of overfitting and reached stationary behaviour at the time of training termination. The model took approximately 73 h to train on the described laptop.

3.2. Segmentation results and discussion

Inference, i.e. segmentation of an entire sample (virtual or physical), took less than 20 s including all data input and output. To evaluate the segmentation performance of the model, three metrics were used. The *Jaccard index*[2] was computed on the testing set (composing synthetic but unseen data) using the same patch size and stride as for the Jaccard loss function during training and validation, but with a batch size of 1. The *Precision* metric P, which represents the number of true positives divided by the sum of the true positives and false negatives, was computed[3] as

$$P(A,B) = \frac{1}{n_\mathrm{c}} \sum_{l=1}^{n_\mathrm{c}} \frac{\sum_{i=1}^{n_x} \sum_{j=1}^{n_y} \sum_{k=1}^{n_z} A_{ijkl} B_{ijkl} + \epsilon}{\sum_{i=1}^{n_x} \sum_{j=1}^{n_y} \sum_{k=1}^{n_z} A_{ijkl}(1 - B_{ijkl}) + \epsilon}, \tag{2}$$

whereas for the defined loss, the output of the network is passed through a softmax function. Finally, the *Recall* metric R, defined as the ratio between true positives and the sum of true positives and false negatives, was computed as

$$R(A,B) = \frac{1}{n_\mathrm{c}} \sum_{l=1}^{n_\mathrm{c}} \frac{\sum_{i=1}^{n_x} \sum_{j=1}^{n_y} \sum_{k=1}^{n_z} A_{ijkl} B_{ijkl} + \epsilon}{\sum_{i=1}^{n_x} \sum_{j=1}^{n_y} \sum_{k=1}^{n_z} (1 - A_{ijkl}) B_{ijkl} + \epsilon}. \tag{3}$$

---

[2] This is computed like in Eq. (1) removing '1−' from the sum over n_c.

[3] using the same notation as in Section 3.1.

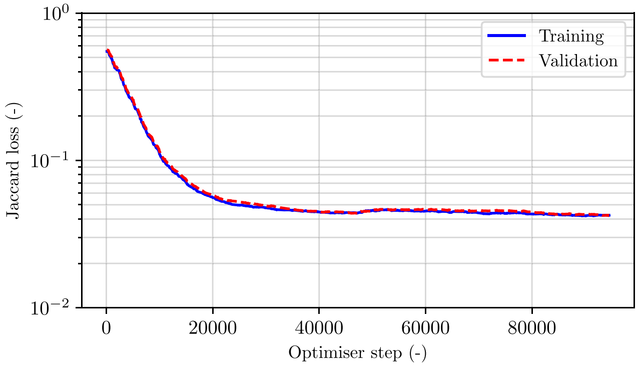


Fig. 3. A lin-log plot of the segmentation model loss is shown. Smoothing according to $y[n+1] = 0.01x[n] + 0.99y[n]$ has been applied to the data, where $y[n]$ is the smoothed data at step n and $x[n]$ is the raw data.

Table 1
Evaluation metrics computed over the testing set. A higher number is better, where 1.0 is the maximum possible value.

Metric	Average	Minimum	Maximum
Jaccard index	0.958	0.897	0.977
Precision	0.978	0.948	0.988
Recall	0.978	0.942	0.988

The ϵ parameter for all metrics uses the same value as in Section 3.1.

The average, the minimum, and the maximum of each metric are summarised in Table 1. It can be seen that the three metrics are very favourable. This indicates that the network has been sufficiently trained. The high minimums show that there is no overfitting that materialises through losing predictability on challenging outliers.

The XRCT scan of the physical sample was segmented[4] with the trained model, using the same patch size, stride and similar normalisation as during training. However,since the glass fibre tracer yarns caused some high contrast artifacts not present in the training data, a $0.6\,\mathrm{cm}^{-1}$ cut off value was used before normalisation (all voxels with a intensity value above 0.6 were set to 0.6). As overlapping patches were used to aid continuity and to prevent sharp borders between patches, the network outputs were summed up and the highest probability material class was selected when reassembling the patches. Refer to Fig. 4 for images of the XRCT data and the corresponding segmentations. The presented segmentations displays highly improved continuity of the weft yarns when compared with previously presented results in [33].

4. Computational RVE homogenisation

Following the segmentation of the XRCT reconstructions the next step in the pipeline is the creation of FE models of the considered RVEs. These models are used to predict the as-manufactured macroscopic elastic properties of the material (with statistical variations) through computational homogenisation of multiple UCs. In this study, the commercial FE code LS-DYNA is used for the numerical analysis, with all FE simulations carried out on an HPC cluster. It is however reiterated, that all novel contributions in this work are open source, and that any reproduction of the study can exchange the particular finite element solver without altering the conclusions.

Generating the RVE models requires, in this case, three distinct steps. Each of these will be discussed in greater detail in the following sections. First, the segmented volume must be discretised into continuum elements, where each element is assigned one of four material classes (matrix, air, warp yarn, or weft yarn). Then, the general input file structure along with keywords for the specific FE-solver must be identified, and the appropriate periodic boundary conditions must be defined. Next, a micromechanical model must be selected to provide the material properties for the yarns. For this model, having the elastic properties of the carbon fibre and matrix as inputs (known from, e.g., technical data sheets), the local fibre volume fraction of the yarn must be approximated element-wise. Finally, the local orientation of the fibre reinforcements within the yarns must be predicted and mapped element-wise as well.

4.1. Finite element model

As previously mentioned, the RVE FE analyses were performed using the commercial code LS-DYNA. More specifically, linear, static implicit analyses were considered. Due to the large number of elements in these models (see more details below), a direct solver is infeasible. Therefore, the iterative conjugate gradient solver with a local symmetric Gauss–Seidel pre-conditioner was used.

When meshing such RVEs, either non-conformal structured meshes or conformal meshes can be considered. Conformal meshes, which explicitly separate and follow the true geometry of the yarns and matrix pockets, have their benefits. However, they require specialised and intricate mesh generation tools. As an alternative non-conformal structured meshes are a favourable alternative, especially since the considered analysis in this work pertain strictly to the elastic regime. It has been demonstrated by Doitrand et al. [47] that the use of non-conformal hexahedral elements is sufficient as long as they are small enough to accurately represent the yarn fibre volume fraction.

One approach, and what is considered in this work, is to perform a one-to-one mapping from each voxel in the segmented RVE volume to an 8-node hexahedral element, with nodes positioned at the corners of the corresponding voxel. Specifically in this case, the RVEs are discretised with LS-DYNA's trilinear and under-integrated `ELEMENT_SOLID_ORTHO` elements (`ELFORM = 1`).

A benefit of the structured mesh is that the application of periodic boundary conditions becomes straightforward. The absolute formulation of the periodic boundary condition equations, as presented by Garoz et al. [48], was used. This formulation uses three corner nodes as so-called 'control nodes' and one fixed 'reference node', making it easy to impose the desired macroscopic infinitesimal strain tensor $\boldsymbol{\varepsilon}$. This is done by prescribing the displacement on the three control nodes (and fixing one node) as

$$\boldsymbol{u}_i = \boldsymbol{\varepsilon} \cdot \boldsymbol{X}_i \quad i \in \text{control node} \quad (\boldsymbol{u}_{\text{fixed}} = \boldsymbol{X}_{\text{fixed}} = 0), \tag{4}$$

where $\boldsymbol{u}_i$ and $\boldsymbol{X}_i$ are the node displacements and positions, respectively.

---

[4] Before segmentation, the reconstructed volume was rotated such that the UC boundaries are aligned with the coordinate axes.

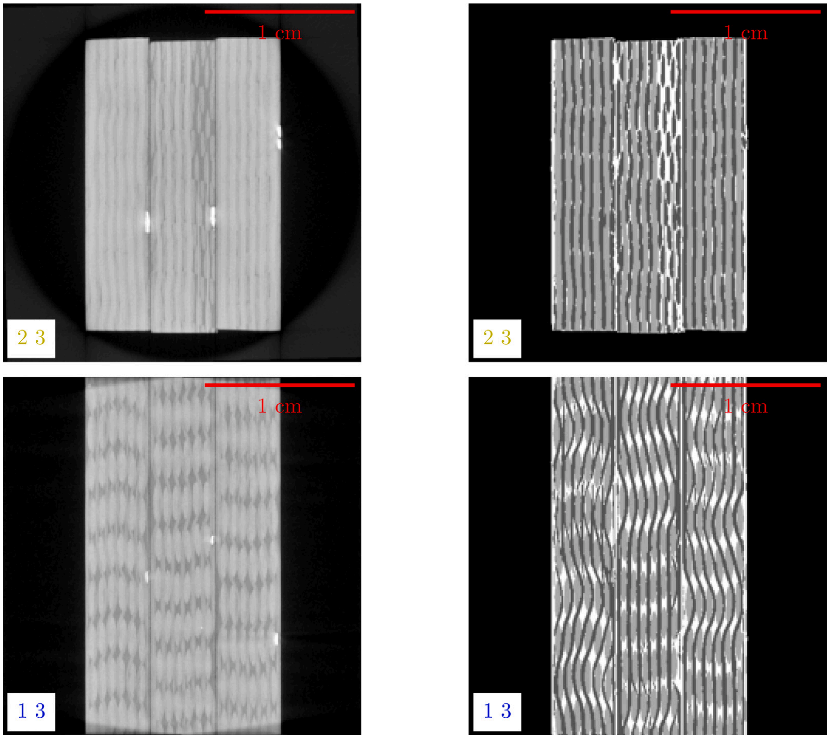


Fig. 4. Two perpendicular slices of the rotated XRCT reconstruction (left) is shown next to their corresponding segmentations (right). For the XRCT scan brighter values correspond to higher attenuation (note the very bright tracer yarns). For the segmentation, black refers to air, dark grey refers to weft yarns, light grey refers to warp yarns, and white refers to pure matrix.

By performing a uniaxial strain controlled load simulation for each of the six strain components, each column of the macroscopic Voigt stiffness matrix can be computed as the resulting macroscopic stress components scaled by the inverse of the applied control strain. In this context, the macroscopic stress tensor, $\bar{\boldsymbol{\sigma}}$, is computed from the stress tensor in each integration point, $\boldsymbol{\sigma}$, as

$$\bar{\boldsymbol{\sigma}} = \frac{1}{V}\int_V \boldsymbol{\sigma}\, dV, \tag{5}$$

where V is the total volume of the RVE. Note that as all elements are equally sized in the voxel mesh, and one integration point per element is used, it is possible to compute the macroscale stress tensor by a simple average over all elements.

The orthotropic linear elastic material model `MAT_ORTHOTROPIC_ELASTIC` is used for all elements. The pure matrix phase is simply assigned isotropic matrix material properties in the orthotropic linear elastic material model. In the case of the impregnated carbon fibre yarns, the transversely isotropic elastic properties are calculated using a micromechanical model that is discussed in greater detail in Section 4.2.

Finally, in this context, it should be noted that LS-DYNA computes output stresses in a finite deformation setting even if the linear momentum balance was solved for infinitesimal strains. To make this discrepancy negligible, a control strain of 0.01% was used for all load cases.

4.2. Micro-mechanical model and constituent properties

Since the FE models resolve individual yarns, but not individual fibres, the yarns are modelled as homogeneous and transversely isotropic. To predict the linear elastic, transversely isotropic properties of the yarns, the Chamis micro mechanical model is used [49]. The model was initially developed for predicting UD laminate properties, but has been successfully used for the yarns of textile composites by several authors, cf. Stig et al. [50], Pierreux et al. [51], and Koptelov et al. [52]. To be specific, the yarn properties ($\bullet^{\mathrm{y}}$) are obtained in terms of transversely isotropic fibre properties ($\bullet^{\mathrm{f}}$), isotropic matrix properties ($\bullet^{\mathrm{m}}$) and the (local) volume fraction of fibres (V^{f}) in each yarn as

$$\begin{aligned}
&E^{\mathrm{y}}_{xx} = V^{\mathrm{f}}E^{\mathrm{f}}_{xx} + (1 - V^{\mathrm{f}})E_{\mathrm{m}} && \nu^{\mathrm{y}}_{xy} = V^{\mathrm{f}}\nu^{\mathrm{f}}_{xy} + (1 - V^{\mathrm{f}})\nu^{\mathrm{m}} = \nu^{\mathrm{y}}_{xz}\\
&E^{\mathrm{y}}_{yy} = \frac{E^{\mathrm{m}}}{1 - \sqrt{V^{\mathrm{f}}}\left(1 - E^{\mathrm{m}}/E^{\mathrm{f}}_{yy}\right)} = E^{\mathrm{y}}_{zz} &&\\
&G^{\mathrm{y}}_{xy} = \frac{G^{\mathrm{m}}}{1 - \sqrt{V^{\mathrm{f}}}\left(1 - G^{\mathrm{m}}/G^{\mathrm{f}}_{xy}\right)} = G^{\mathrm{y}}_{xz} &&\\
&G^{\mathrm{y}}_{yz} = \frac{G^{\mathrm{m}}}{1 - \sqrt{V^{\mathrm{f}}}\left(1 - G^{\mathrm{m}}/G^{\mathrm{f}}_{yz}\right)} && \nu^{\mathrm{y}}_{yz} = \frac{E^{\mathrm{y}}_{yy}}{2G^{\mathrm{y}}_{yz}} - 1
\end{aligned} \tag{6}$$

Here x denotes the direction parallel to the yarn (or fibre), while y and z denotes the two transverse directions.

The calculation of the yarn fibre volume fraction is discussed in more detail in Section 4.3. The constituent mechanical properties used in this work are given in Table 2. For the fibres, the longitudinal modulus is taken from the manufacturer's data sheet [37], and the remaining properties are taken from the third world-wide failure exercise [53]. For the matrix, the (tensile) modulus of RTM6 is also taken

Table 2
The material constituent mechanical properties.

Property	Value	Reference
$E_{xx\mathrm{f}}$	276 GPa	[37,53]
$E_{yy\mathrm{f}}$	19 GPa	[53]
$G_{xy\mathrm{f}}$	27 GPa	[53]
$G_{yz\mathrm{f}}$	7 GPa	[53]
$\nu_{xy\mathrm{f}}$	0.2	[53]
E_{m}	3.1 GPa	[38,54]
ν_{m}	0.36	[54]

from the manufacturer's data sheet [38], which aligns well to what was (independently) measured by Exner et al. [54]. Further, Exner et al. [54] also reported the Poisson's ratio of the RTM6, which was used in this work.

4.3. Fibre orientation mapping

As the next step, the local yarn orientations need to be estimated and assigned to each element. To determine the local fibre orientation it is assumed that the fibres can be estimated to have no in-plane misalignment. Based on this assumption, and speaking broadly, the local fibre orientation is inferred with a structure tensor analysis of the segmented RVE, carried out in 2D on a per-slice basis. The Python package `structure-tensor` [55] was used for this analysis. For both yarn types respectively, the volume was sliced along the axis that is perpendicular to both the yarn direction and the out-of-plane direction. An example is visualised in Fig. 5.

The binary nature of the segmentation yields little gradient information inside the yarns. This makes the structure tensor difficult to apply directly. Therefore, the Euclidean distance transform was computed for each slice, in which each segmented yarn voxel in the slice is assigned its distance to the closest voxel of any other class in the same slice. This creates ridge-shaped structures that follow the yarns, creating a large gradient perpendicular to the yarns and a near zero gradient along the yarns, which improves the accuracy of the structure tensor analysis, cf. Fig. 5a.

Furthermore, numerical issues arise when performing structure tensor analysis of large regions of constant value. As a remedy, uniform random noise between 0 and 0.1 (the higher value corresponding to a 10th of a voxel) was added to the distance value of each pixel. Thereafter, the image was smoothed with a Gaussian filter with kernel size 1. The noised, smoothed, distance transformed slice, cf. Fig. 5c, was analysed with a ρ parameter (integration kernel size) of 2.0 and a σ (gradient kernel size) parameter of 0.25.

Finally, the eigenvectors resulting from the structure tensor analysis were examined. In the structure tensor analysis the eigenvector associated with the smallest eigenvalue of the structure tensor represents the direction of minimal variation in the underlying image [56], in this case corresponding to the local orientation of the yarns. As an example, considering an image with stripes. For this case, the gradient is large when going from a stripe to background, while small when following the stripes.

To visualise the outcome of the structure tensor analysis, the weft yarn orientations of one UC are shown in Fig. 6. Analysing the direction of eigenvectors, it can be found that the level of crimp (the magnitude of the out-of-plane component of the fibre orientation vectors) in the weft is very consistent, even though it was not strictly enforced. The average weft yarn unit orientation vector out-of-plane component magnitude was determined to be 0.135 with a standard deviation of only 0.003. In addition, the warp yarns display very low crimp in all the identified UCs and are not visualised. The average warp out-of-plane component magnitude was found to be 0.047 with a standard deviation of 0.003.

4.4. Fibre volume fraction mapping

For the purpose of computing the local fibre volume fraction of the yarns, it should be noted that in the case of the warp yarns, the segmentation is able to clearly differentiate between one yarn instance and the next throughout the volume. This, in simple terms, would make the calculation of the local yarn fibre volume fraction straightforward given the total area of the yarn, the number of filaments in the yarn and the diameter of each filament. One challenging aspect to this material, however, is that in some slices the weft yarns are highly compacted and appear to merge with one another, cf. the side pieces in the 2–3 plane in Fig. 4. This makes it difficult to differentiate one weft yarn instance from the next. Therefore, to estimate the local fibre volume fraction in each element of the yarns the a priori known number of yarns and the number of fibres per yarn was exploited.

It is assumed that, for each yarn type, the fibre volume fraction can be approximated as constant per cross section of the segmented volume along each respective yarn weave direction. As such, the volume fraction of each yarn type is estimated by slicing the segmented UC along the axis parallel to the respective yarn direction. Refer to Fig. 7 for a representative schematic. A first estimate of the slice-wise average volume fraction is given by the ratio between the theoretical total area of the fibres in a yarn type in a cross section and the area covered by voxels segmented as the same yarn type in the same cross section. The theoretical area covered by the fibres in a cross section is computed as

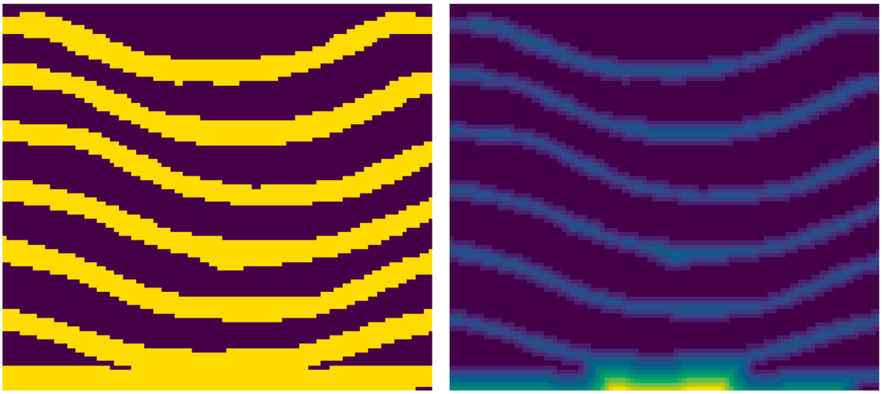
$$A_{\mathrm{fibre}} = N_{\mathrm{fibre}} N_{\mathrm{yarn}} \frac{\pi d_{\mathrm{fibre}}^2}{4}, \tag{7}$$

where d_{fibre} is the fibre diameter, and N_{fibre} and N_{yarn} are the number of fibres per yarn and number of yarns per cross section, respectively (which depend on the yarn type). This first estimate is accurate if the yarns are perpendicular to the cross section in question. In reality, however, the yarns are not straight. This was compensated for by scaling the area of each voxel with the unit orientation vector component in the yarn direction. That is, voxels where yarns are perpendicular to the cross section received a scale factor of 1, while other voxels representing up-wards or down-wards crimping yarn segments obtained a scale factor less than 1. In summary, the fibre volume fraction was thus estimated slice-wise by computing the ratio between the theoretical area covered by the fibres and the orientation-compensated area of the segmented voxels.

Finally, it is noted that, to avoid generating too many material cards, the volume fractions are binned into 20 bins that are equally spaced between the minimum yarn fibre volume fraction and the maximum yarn fibre volume fraction. For illustrative purposed, the fibre volume fraction inside both the weft and warp yarns is displayed in Fig. 8 for the same UC shown previously. As expected, the fibre volume fraction of the weft yarns is highest as it is squeezed between the rows of warp yarns. Further, it should be noted that the yarn fibre volume fractions vary between 64%–78%.

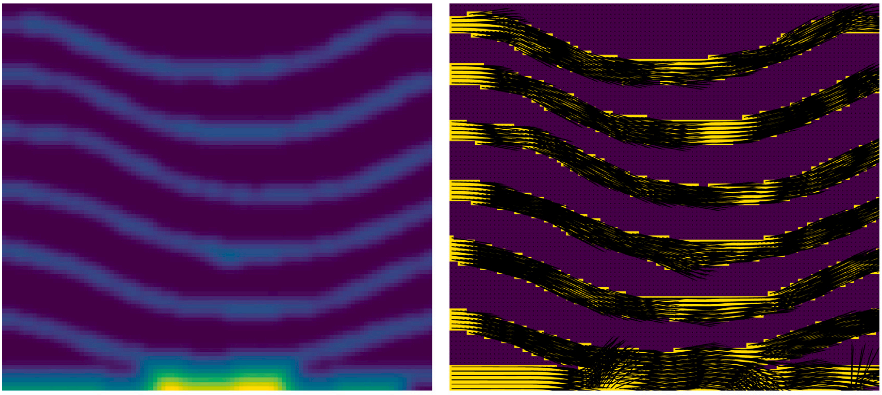
4.5. Results, discussion and comparison to experimental results

From the segmented data, six non-overlapping UCs were identified by hand. The UCs were selected where the segmentation was qualitatively the most accurate. It is here emphasised that this was the only manual labour performed in the entire pipeline. After the initial selection, the described segmentation, meshing and material mapping procedures were performed on all six UCs, respectively. Separate FE analyses of each of the identified UCs were then performed to estimate the elastic properties and their statistics. As previously discussed, this involves the application of six distinct load cases using periodic boundary conditions, each of which represents a non-zero strain component. By carrying out first order computational homogenisation the entire stiffness tensor of the homogenised material sample can be computed.



(a) The weft yarns are yellow (foreground) while all other phases are purple (background).

(b) The Euclidean distance transform of the image in (a). A bright pixel corresponds to a larger distance value.



(c) Noise has been added to the image in (b) and smoothed with a Gaussian kernel.

(d) The eigen vectors resulting from a structure tensor analysis of the image in (c).

Fig. 5. The orientation analysis of a weft yarn is demonstrated. A small section of a slice along the warp direction is shown.

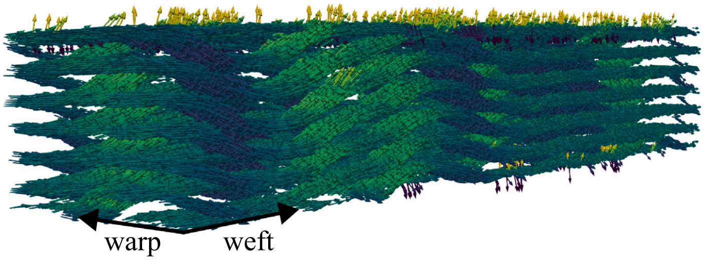


Fig. 6. The weft yarn orientation vectors in one of the identified UCs are visualised. The colouring aids visualisation contrast, were bright colours represents a positive out-of-plane component and dark represents negative.

For each of the 6 UCs, 3 tensile (along weft, along warp, and out-of-plane), and 3 pure shear (in-plane, weft out-of-plane, and warp out-of-plane) tests were performed. For each load case, the macroscopic stress scaled by the macroscopic strain resulted in one column of the Voigt stiffness matrix derived for the corresponding UC. Fig. 9 shows the normal stress in the weft direction, for the load case of pure elongation in the weft direction. The figure concerns the UC (no. 1) that is displayed in Figs. 6 and 8. The values of the homogenised stiffness

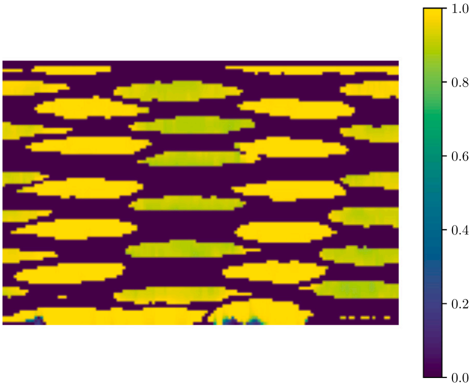


Fig. 7. The weft cross sections are shown in a slice along the weft direction. The colour bar refers to the voxel area scale factor due to the fibre direction.

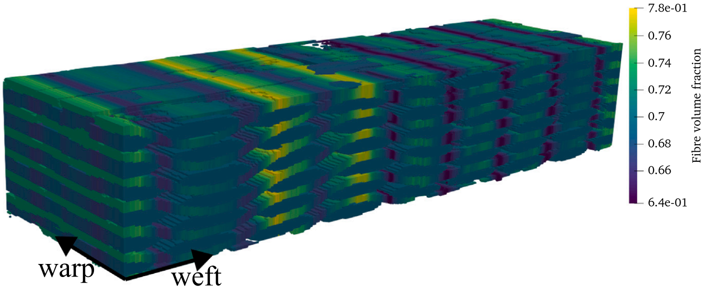


Fig. 8. The fibre volume fractions of one of the identified UCs are displayed.

Table 3
The computed homogenised mean and standard deviation of the macroscopic engineering tensile and shear moduli for the six UCs are shown together with the measured in-plane experimental values. The values are given in GPa.

Parameter	E_{11}	E_{22}	E_{33}	G_{12}	G_{23}	G_{13}
Experiment [6]	68.65	61.43		3.91		
Sim. mean	71.92	57.75	9.10	4.11	3.60	3.28
Sim. std dev	0.62	1.11	0.04	0.08	0.08	0.06

tensor of the same UC are given as

$$\mathbf{C}_1 = \begin{bmatrix} 72.9544 & 5.1803 & 4.1161 & 0.0155 & 0.0014 & -0.3598 \\ 5.1798 & 60.1845 & 4.9221 & -0.0947 & -0.0275 & -0.0392 \\ 4.1161 & 4.9221 & 9.6476 & 0.0023 & 0.0042 & -0.0158 \\ 0.0046 & -0.0228 & 0.0007 & 4.1954 & -0.0066 & -0.0008 \\ 0.0004 & -0.0066 & 0.0012 & -0.0066 & 3.5235 & -0.0006 \\ -0.0896 & -0.0097 & -0.0038 & -0.0008 & -0.0006 & 3.2249 \end{bmatrix} \text{GPa}, \tag{8}$$

where $\mathbf{C}_1$ is the Voigt matrix representation thereof. It is evident that the resulting stiffness tensor demonstrates slight asymmetry. However, this is not unexpected, and given the relative magnitude of stiffness components, it has still been deemed acceptable. In addition, the lower left and upper right quadrants of the Voigt matrix representation of the stiffness tensor are not exactly zero as required by orthotropy. It is however deemed that, since the values in these quadrants are orders of magnitude smaller than the true non-zero components, the assumption of elastic orthotropy is still reasonable.

Table 4
The computed mean and standard deviation of the homogenised macroscopic Poisson's ratios for the six UCs.

Parameter	ν_{12}	ν_{21}	ν_{23}	ν_{32}	ν_{13}	ν_{31}
Mean	0.0509	0.0409	0.4922	0.0776	0.3954	0.0500
Std dev	0.0031	0.0034	0.0034	0.0019	0.0032	0.0005

To compensate for the minor asymmetry of the stiffness tensor, the extraction of engineering macroscopic properties was performed from the symmetrised Voigt stiffness matrix given by

$$\mathbf{C}_{\text{sym}} = \frac{1}{2}\left(\mathbf{C} + \mathbf{C}^{T}\right). \tag{9}$$

Thereafter, the nine macroscopic elastic engineering parameters of orthotropy (using the coordinate convention in Fig. 1(a)) are easily identified from the compliance Voigt matrix through

$$\mathbf{S}_{\text{sym}} = \mathbf{C}_{\text{sym}}^{-1} = \begin{bmatrix} \frac{1}{E_{11}} & -\frac{\nu_{21}}{E_{22}} & -\frac{\nu_{31}}{E_{33}} & 0 & 0 & 0 \\ -\frac{\nu_{12}}{E_{11}} & \frac{1}{E_{22}} & -\frac{\nu_{32}}{E_{33}} & 0 & 0 & 0 \\ -\frac{\nu_{13}}{E_{11}} & -\frac{\nu_{23}}{E_{22}} & \frac{1}{E_{33}} & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{G_{12}} & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{G_{23}} & 0 \\ 0 & 0 & 0 & 0 & 0 & \frac{1}{G_{13}} \end{bmatrix}. \tag{10}$$

The macroscopic elastic parameters, computed on all of the analysed UCs, are displayed in Fig. 10. The average values of the tensile and shear moduli, and their standard deviations, are shown in Table 3. For comparison, also experimentally obtained tensile and shear moduli

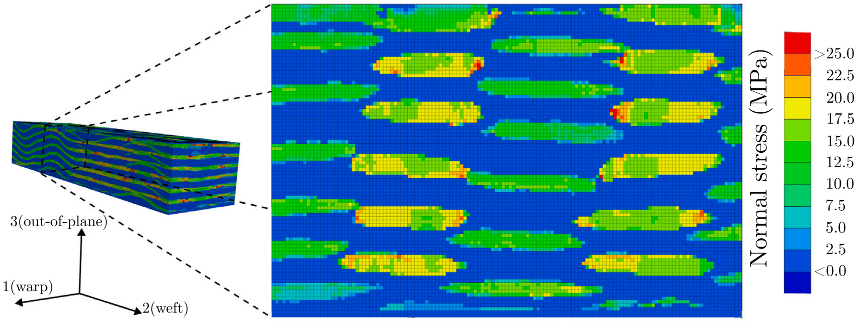


Fig. 9. The normal stress in the weft (2) direction is shown for the pure elongation in the weft direction load case. Note the one-to-one correspondence between elements and segmentation voxels. The figure illustrates how the weft yarns carry the stress.

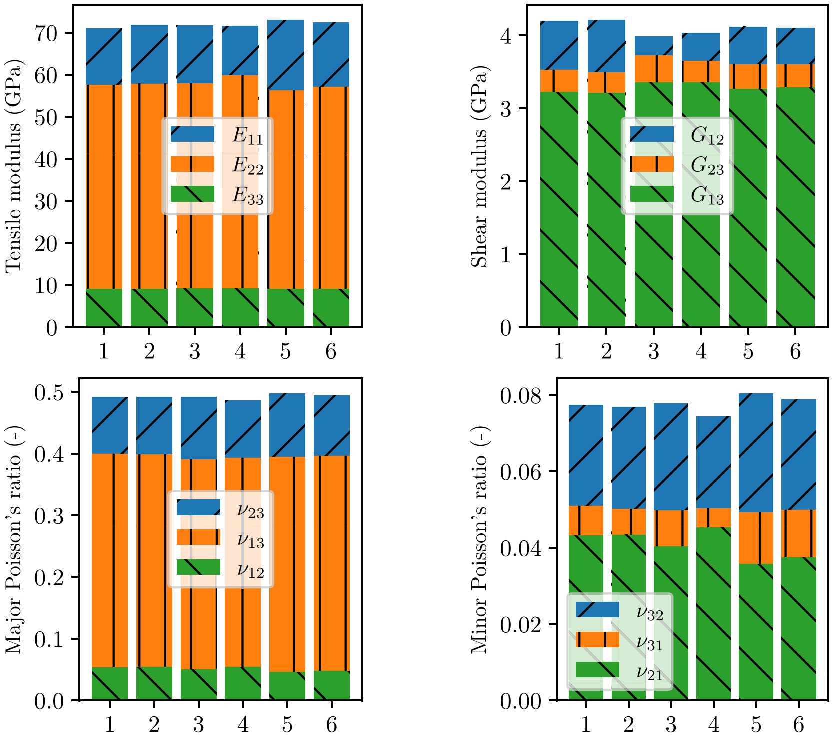


Fig. 10. The macroscopic elastic properties for each UC are displayed.

are presented in the same table. These moduli have been extracted for the exact same material from the cyclic test data obtained by Oddy et al. [6]. More specifically, each separate modulus has been determined via a linear least square fit of the stress–strain data of the earliest possible load cycle including the full strain range $\Delta\varepsilon = \varepsilon_{max} - \varepsilon_{min} = 0.003 - 0.001$[5] (as stipulated by *ASTM D3039* [57]). Furthermore, the average Poisson's ratios and their standard deviations are shown in Table 4.

From these results, it is observed that the predicted average of the tensile modulus in the warp direction differs with less than 4.7% compared to the experimental value. Similarly, the predicted average tensile modulus in the weft direction and the predicted average shear modulus show deviations of less than 6.4% and around 5%, respectively.

The small deviations from experimental values demonstrate the capability of the proposed pipeline in predicting the as-manufactured elastic properties of a 3D-woven CFRP. Moreover, owing to the fact

---

[5] The test data in [6] was obtained through cyclic loading with a continuously increasing strain range for each cycle. The 'earliest possible load cycle' is thereby the first load cycle that contains the whole strain range

that several UCs can be investigated from only one XRCT scan, the variability in a material can be easily investigated.

It was surprising that the amount of out-of-plane fibre misalignment was as consistent as observed. At the same time, this explains the consistency in the predicted elastic properties. Still, as seen in the bottom left part of Fig. 4, there is a noticeable difference in warp yarn configuration between the three stacked pieces. The two pieces on each side have their warp yarns arranged in a U-shape, while the centre piece displays more straightly stacked warp yarns. It is likely that these morphological differences would have a larger impact if deformations increase (requiring finite strain simulations). As Oddy et al. [6] reported from their experiments, the material under study displayed yarn realignment for large strains. It is, however, evident that to extract linearly elastic properties (relevant for small strains), as long as the fibre volume fraction and orientation distribution is captured accurately enough, the analysis of a single UC is enough. This indicates that the samples under study have been manufactured in a fairly consistent manner.

5. Conclusion and future outlook

In this work, a fully automated and open source pipeline for the prediction of the elastic properties of 3D-woven composites in their as-manufactured state has been presented. The pipeline allows for automated segmentation of challenging large field of view scans, automated meshing and assignment of material orientations and properties as well as automated definitions of periodic boundary conditions for computational homogenisation. It includes a novel method for estimating the local yarn orientation in cases where instance segmentation is challenging or even impossible. In more detail, the 3D U-net architecture showed great promise in segmenting large field of view XRCT scans of CFRP composites.

Applying this pipeline to a resin-infused layer-to-layer angle interlock weave, it has been demonstrated that macroscopic elastic properties can be predicted in good agreement with experimental observations. Highly accurate predictions (less than 6.5% deviation from experiments) of the full 3D elastic properties were achieved. This indicates that the pipeline can be an efficient tool for predicting the elastic properties directly from large field of view XRCT scans of 3D-woven composites.

The execution of the pipeline is computationally efficient. When the segmentation model has been pre-trained, the time from a reconstructed volume to an FE input file is less than a minute and fully automated. Being automated and efficient, the pipeline allows for several UCs to be generated from a single scan, which unlocks the possibility to estimate statistical information of the investigated material properties. Interestingly, however, for the material under study, it was found that the elastic properties had very little statistical variation between UCs, even though the meso scale geometry had noticeable differences. Future works should investigate how large the variation in the meso scale geometry needs to be to yield significant variation in elastic properties.

Having validated the ability of the pipeline to accurately predict the elastic properties of as-manufactured samples, a relevant follow-up is the investigation of large deformations, plasticity and damage. In order to expand the analysis to these domains, the use of a voxel mesh becomes infeasible due the very large number of elements. Furthermore, a voxel mesh may yield artificial stress concentrations [58] that can erroneously initiate plasticity and damage [59]. As a consequence, future works should incorporate a conformal meshing approach. A meso scale model with conformal meshing of the yarns would however require instance segmentation of individual yarns. This has been done by Blusseau et al. [30], Mendoza et al. [29], and Fourrier et al. [22] among others. Future work could take inspiration from these approaches and generalise to allow for segmenting very low resolution, large field of view scans studied in this work, and to eliminate remaining steps where manual input is necessary.

It is a challenging undertaking to generalise conformal meshing procedures to many large UCs. A possible alternative could be to encode not only phase ground truths in the training data, but also encode the yarn orientations and phase boundaries. Additionally, instead of predicting the segmentation as an intermediate step it would be interesting to investigate inferring geometry descriptions directly in the neural network model. This would require finding an encoding for the geometry information that can handle arbitrary number of yarns, and arbitrary yarn shapes.

CRediT authorship contribution statement

Johan Friemann: Writing – original draft, Validation, Software, Methodology, Investigation, Formal analysis, Data curation. **Carolyn Oddy:** Writing – original draft, Supervision. **Lars P. Mikkelsen:** Writing – review & editing, Supervision, Resources. **Martin Fagerström:** Writing – review & editing, Supervision, Resources, Funding acquisition, 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

The authors acknowledge funding from Horizon Europe through the MSCA Doctoral Network RELIANCE: Reliance: REaL-tIme characterisation of ANisotropic Carbon-based tEchnological fibres, films and composites, grant no. 101073040. The X-ray scans were captured at DTU Center for Advanced Structural and Material Testing (CASMaT) located at the 3D imaging Centre (3DIM). The FE simulations were performed on resources at the Chalmers Centre for Computational Science and Engineering (C3SE) provided by the Swedish National Infrastructure for Computing (SNIC). The authors want to thank Dr. Lei Liu for help with the formulation of periodic boundary conditions.

Data availability

All data not explicitly referred to in the manuscript will be made available on request.

References

[1] K. Jing, S. Xie, Y. Zhang, H. Zhou, H. Yan, Impact resistance of 3D woven fabrics and composites: A review, Thin-Walled Struct. 213 (2025) 113262.

[2] M.N. Saleh, C. Soutis, Recent advancements in mechanical characterisation of 3D woven composites, Mech. Adv. Mater. Mod. Process. 3 (1) (2017).

[3] P. Tan, L. Tong, G. Steven, T. Ishikawa, Behavior of 3D orthogonal woven CFRP composites. Part I. Experimental investigation, Compos. Part A: Appl. Sci. Manuf. 31 (3) (2000) 259–271.

[4] Y. Mahadik, S. Hallett, Effect of fabric compaction and yarn waviness on 3D woven composite compressive properties, Compos. Part A: Appl. Sci. Manuf. 42 (11) (2011) 1592–1600.

[5] K. Leong, B. Lee, I. Herszberg, M. Bannister, The effect of binder path on the tensile properties and failure of multilayer woven CFRP composites, Compos. Sci. Technol. 60 (1) (2000) 149–156.

[6] C. Oddy, M.y. Song, C. Stewart, B. El Said, M. Ekh, S.R. Hallett, M. Fagerström, Cyclic behaviour of 3D-woven composites in tension: Experimental testing and macroscale modelling, Compos. Part A: Appl. Sci. Manuf. 187 (2024) 108354.

[7] L.P. Brown, A.C. Long, Modeling the geometry of textile reinforcements for composites: TexGen, in: P. Boisse (Ed.), Composite Reinforcements for Optimum Performance, second ed., in: Woodhead Publishing Series in Composites Science and Engineering, Woodhead Publishing, 2020, pp. 237–265.

[8] S.V. Lomov, Modeling the geometry of textile composite reinforcements: WiseTex, Composite Reinforcements for Optimum Performance, second ed., Elsevier Ltd., 2020, pp. 199–236.

[9] Y. Wang, X. Sun, Digital-element simulation of textile processes, Compos. Sci. Technol. 61 (2) (2001) 311–319.

[10] G. Zhou, X. Sun, Y. Wang, Multi-chain digital element analysis in textile mechanics, Compos. Sci. Technol. 64 (2) (2004) 239–244.

[11] B.S. El Said, S. Green, S.R. Hallett, Kinematic modelling of 3D woven fabric deformation for structural scale features, Compos. Part A: Appl. Sci. Manuf. 57 (2014) 95–107.

[12] D. Durville, I. Baydoun, H. Moustacas, G. Périé, Y. Wielhorski, Determining the initial configuration and characterizing the mechanical properties of 3D angle-interlock fabrics using finite element simulation, Int. J. Solids Struct. 154 (2018) 97–103.

[13] Z. Yang, Y. Jiao, J. Xie, L. Chen, W. Jiao, X. Li, M. Zhu, Modeling of 3D woven fibre structures by numerical simulation of the weaving process, Compos. Sci. Technol. 206 (2021) 108679.

[14] T. Shao, S. Zhang, W. Liu, R. Liu, W. Xu, High-fidelity architecture modeling and compressive strength prediction of 3D woven composite material, Compos. Struct. 354 (2025) 118775.

[15] Z. Gao, L. Chen, S. Zhao, Z. Zhang, A deformation procedure using position-based dynamics to optimize the geometric model of woven fabrics, Appl. Math. Model. 137 (2025) 115650.

[16] A. Ewert, B. Drach, K. Vasylevskyi, I. Tsukrov, Predicting the overall response of an orthogonal 3D woven composite using simulated and tomography-derived geometry, Compos. Struct. 243 (March) (2020) 112169.

[17] S.U. Lee, S. Chung, R.H. Park, A comparative performance study of several global thresholding techniques for segmentation, Comput. Vis. Graph. Image Process. 52 (2) (1990) 171–190.

[18] J.B. Roerdink, A. Meijster, The watershed transform: Definitions, algorithms and parallelization strategies, Fund. Inform. 41 (1, 2) (2000) 187–228.

[19] R.M. Auenhammer, L.P. Mikkelsen, L.E. Asp, B.J. Blinzler, Automated X-ray computer tomography segmentation method for finite element analysis of non-crimp fabric reinforced composites, Compos. Struct. 256 (August 2020) (2021) 113136.

[20] N. Naouar, E. Vidal-Sallé, J. Schneider, É. Maire, P. Boisse, Meso-scale FE analyses of textile composite reinforcement deformation based on X-ray computed tomography, Compos. Struct. 116 (1) (2014) 165–176.

[21] R.M. Auenhammer, N. Jeppesen, L.P. Mikkelsen, V.A. Dahl, B.J. Blinzler, L.E. Asp, Robust numerical analysis of fibrous composites from X-ray computed tomography image data enabling low resolutions, Compos. Sci. Technol. 224 (2022) 109458.

[22] G. Fourrier, A. Rassineux, F.H. Leroy, M. Hirsekorn, C. Fagiano, E. Baranger, Automated conformal mesh generation chain for woven composites based on CT-scan images with low contrasts, Compos. Struct. 308 (August 2022) (2023) 116673.

[23] J. Bénézech, G. Couégnat, Variational segmentation of textile composite preforms from X-ray computed tomography, Compos. Struct. 230 (August) (2019) 111496.

[24] Y. Sinchuk, P. Kibleur, J. Aelterman, M.N. Boone, W. Van Paepegem, Variational and deep learning segmentation of very-low-contrast X-ray computed tomography images of carbon/epoxy woven composites, Materials 13 (4) (2020).

[25] V. Pidou-Brion, Y. Le Guilloux, Active yarn meshes for segmentation on X-ray computed tomography of textile composite materials at the mesoscopic scale, Compos. Struct. 281 (December 2021) (2022).

[26] Y. Pannier, P. Coupé, T. Garrigues, M. Gueguen, P. Carré, Automatic segmentation and fibre orientation estimation from low resolution X-ray computed tomography images of 3D woven composites, Compos. Struct. 318 (April) (2023) 117087.

[27] C. Tang, J. Zou, Y. Xiong, B. Liang, W. Zhang, Automatic reconstruction of closely packed fabric composite RVEs using yarn-level micro-CT images processed by convolutional neural networks (CNNs) and based on physical characteristics, Compos. Sci. Technol. 252 (April) (2024) 110616.

[28] Z. Ying, H. Chen, X. Pan, Z. Wu, L. Peng, X. Cheng, J. Ou, Geometric reconstruction of mesoscopic models for three-dimensional woven composites using micro-CT technology, Compos. Struct. 371 (2025) 119466.

[29] A. Mendoza, R. Trullo, Y. Wielhorski, Descriptive modeling of textiles using FE simulations and deep learning, Compos. Sci. Technol. 213 (2021) 108897.

[30] S. Blusseau, Y. Wielhorski, Z. Haddad, S. Velasco-Forero, Instance segmentation of 3D woven fabric from tomography images by Deep Learning and morphological pseudo-labeling, Compos. Part B: Eng. 247 (March) (2022).

[31] K. Zheng, C. Wu, H. Chen, X. Zhang, Z. Wang, Z. Pan, B. Qiu, Z. Wu, Improved 3D image segmentation for X-ray tomographic data of biaxial warp-knitted composites, J. Reinf. Plast. Compos. 44 (7–8) (2023) 344–357.

[32] A. Tsamos, S. Evsevleev, R. Fioresi, F. Faglioni, G. Bruno, Synthetic Data Generation for Automatic Segmentation of X-ray Computed Tomography Reconstructions of Complex Microstructures, J. Imaging 9 (2) (2023).

[33] J. Friemann, L.P. Mikkelsen, C. Oddy, M. Fagerström, Synthetic, automatically labelled training data for machine learning based X-ray CT image segmentation: Application to 3D-textile carbon fibre reinforced composites, Compos. Part B: Eng. 305 (2025) 112656.

[34] J. Friemann, Textomos [software], 2025, URL: https://github.com/Johan-Friemann/textomos.

[35] K. Jing, H. Zhou, H. Wang, H. Yan, S. Xie, Prediction and adjustment of elastic properties in three-dimensional orthogonal woven composites through dual-scale modeling, Polym. Compos. 45 (6) (2024) 5430–5448.

[36] A. Ahmed, I. Ud Din, R. Ramachandran, M. Sikandar Bathusha, W. Cantwell, K.A. Khan, Micro-CT driven characterization and validation of constituents' properties used in origami-inspired foldable composites fabricated via different manufacturing processes, Compos. Struct. 322 (2023) 117382.

[37] Hexcell, HexTow® IM7 Carbon Fiber, 2023, pp. 1–3, Product data sheet, https://www.hexcel.com/wp-content/uploads/2025/12/IM7_HexTow_DataSheet.pdf.

[38] Hexcell, HexFlow® RTM6/RTM6-2 180 °C Epoxy System for Resin Transfer Molding and Infusion Technologies, 2023, pp. 1–3, Product data sheet, https://www.hexcel.com/wp-content/uploads/2025/12/RTM6_RTM62_HexFlow_DataSheet.pdf.

[39] L.A. Feldkamp, L.C. Davis, J.W. Kress, Practical cone-beam algorithm, J. Opt. Soc. Amer. A 1 (6) (1984) 612–619.

[40] W. van Aarle, W.J. Palenstijn, J. Cant, E. Janssens, F. Bleichrodt, A. Dabravolski, J. De Beenhouwer, K.J. Batenburg, J. Sijbers, Fast and flexible X-ray tomography using the ASTRA toolbox, Opt. Express 24 (22) (2016) 25129–25147.

[41] J. Friemann, X-ray CT scan of layer to layer angle interlock composite sample with hand segmentation [dataset], 2025, URL: https://zenodo.org/records/14891845.

[42] Y. Sinchuk, Y. Wielhorski, A. Mendoza, S. Blusseau, S. Velasco-Forero, Automatic yarn path extraction of large 3D interlock woven fabrics with confidence estimation, Compos. Part A: Appl. Sci. Manuf. 186 (2024) 108396.

[43] Y. Cao, Z. Nie, F. Sun, Q. Zhao, X. Yang, C. Li, FabricFlow: Automatic segmentation of 3D closely packed woven fabric in low-contrast CT images through gradient flow tracking, Mater. Today Commun. 48 (2025) 113294.

[44] A. Paszke, S. Gross, F. Massa, A. Lerer, J. Bradbury, G. Chanan, T. Killeen, Z. Lin, N. Gimelshein, L. Antiga, A. Desmaison, A. Kopf, E. Yang, Z. DeVito, M. Raison, A. Tejani, S. Chilamkurthy, B. Steiner, L. Fang, J. Bai, S. Chintala, Pytorch: An imperative style, high-performance deep learning library, in: Advances in Neural Information Processing Systems 32, Curran Associates, Inc., 2019, pp. 8024–8035, URL: http://papers.neurips.cc/paper/9015-pytorch-an-imperative-style-high-performance-deep-learning-library.pdf.

[45] I. Loshchilov, F. Hutter, SGDR: Stochastic gradient descent with warm restarts, 2016, URL: https://arxiv.org/abs/1608.03983.

[46] Y. Liu, Y. Ge, R. Pan, A. Kang, T. Zhang, Theoretical analysis on how learning rate warmup accelerates convergence, 2025, URL: https://arxiv.org/abs/2509.07972.

[47] A. Doitrand, C. Fagiano, F.-X. Irisarri, M. Hirsekorn, Comparison between voxel and consistent meso-scale models of woven composites, Compos. Part A: Appl. Sci. Manuf. 73 (2015) 143–154.

[48] D. Garoz, F. Gilabert, R. Sevenois, S. Spronk, W. Van Paepegem, Consistent application of periodic boundary conditions in implicit and explicit finite element simulations of damage in composites, Compos. Part B: Eng. 168 (2019) 254–266.

[49] C.C. Chamis, Mechanics of composite materials: Past, present, and future, J. Compos. Technol. Res. 11 (1) (1989) 3.

[50] F. Stig, S. Hallström, A modelling framework for composites containing 3D reinforcement, Compos. Struct. 94 (9) (2012) 2895–2901.

[51] G. Pierreux, D. Van Hemelrijck, T.J. Massart, Automated generation of 3D orthogonal woven composites RVEs including yarn cross-section variations, Compos. Sci. Technol. 176 (2019) 90–102.

[52] A. Koptelov, A. Thompson, S.R. Hallett, B. El Said, A deep learning approach for predicting the architecture of 3D textile fabrics, Mater. Des. 239 (2024) 112803.

[53] A. Kaddour, M. Hinton, P. Smith, S. Li, Mechanical properties and details of composite laminates for the test cases used in the third world-wide failure exercise, J. Compos. Mater. 47 (20–21) (2013) 2427–2442.

[54] W. Exner, R. Hein, T. Mahrholz, P. Wierach, H.P. Monner, M. Sinapius, Impact of nanoparticles on the process-induced distortions of carbon fiber reinforced plastics: An experimental and simulative approach, J. Appl. Polym. Sci. 136 (5) (2018).

[55] N. Jeppesen, L. Mikkelsen, A. Dahl, A. Christensen, V. Dahl, Quantifying effects of manufacturing methods on fiber orientation in unidirectional composites using structure tensor analysis, Compos. Part A: Appl. Sci. Manuf. 149 (2021) 106541.

[56] I. Straumit, S.V. Lomov, M. Wevers, Quantification of the internal structure and automatic generation of voxel models of textile composites from X-ray computed tomography data, Compos. Part A: Appl. Sci. Manuf. 69 (2015) 150–158.

[57] A. International, ASTM D3039 Standard Test Method for Tensile Properties of Polymer Matrix Composite Materials, 2025, http://dx.doi.org/10.1520/d3039_d3039m-08.

[58] G. Fang, B. El Said, D. Ivanov, S.R. Hallett, Smoothing artificial stress concentrations in voxel-based models of textile composites, Compos. Part A: Appl. Sci. Manuf. 80 (2016) 270–284.

[59] T. Zheng, L. Guo, Z. Tang, T. Wang, Z. Li, Comparison of progressive damage simulation of 3D woven composites between voxel and conformal discretization models, Mech. Mater. 158 (2021) 103860.