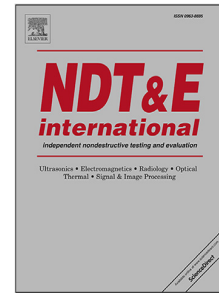


Journal Pre-proof

Novel thresholding method and convolutional neural network for fiber volume content determination from 3D μ CT images

Juliane Blarr, Philipp Kunze, Noah Kresin, Wilfried V. Liebig,
Kaan Inal, Kay A. Weidenmann



PII: S0963-8695(24)00032-X
DOI: <https://doi.org/10.1016/j.ndteint.2024.103067>
Reference: JNDT 103067

To appear in: *NDT and E International*

Received date: 23 August 2023
Revised date: 31 January 2024
Accepted date: 3 February 2024

Please cite this article as: J. Blarr, P. Kunze, N. Kresin et al., Novel thresholding method and convolutional neural network for fiber volume content determination from 3D μ CT images. *NDT and E International* (2024), doi: <https://doi.org/10.1016/j.ndteint.2024.103067>.

This is a PDF file of an article that has undergone enhancements after acceptance, such as the addition of a cover page and metadata, and formatting for readability, but it is not yet the definitive version of record. This version will undergo additional copyediting, typesetting and review before it is published in its final form, but we are providing this version to give early visibility of the article. Please note that, during the production process, errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

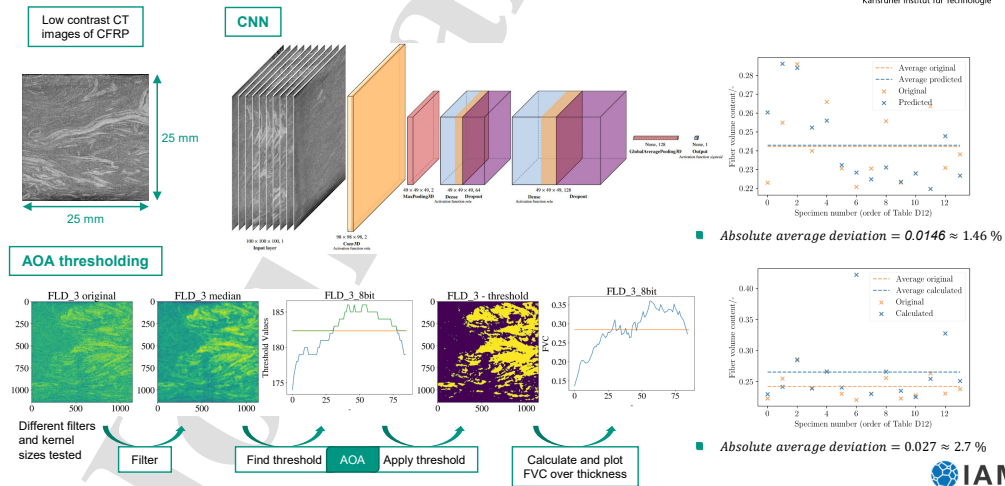
© 2024 Published by Elsevier Ltd.

Graphical Abstract

Novel thresholding method and convolutional neural network for fiber volume content determination from 3D μ CT images

Blarr, Juliane, Kunze, Philipp, Kresin, Noah, Liebig, Wilfried V., Inal, Kaan, Weidenmann, Kay A.

Introducing a novel computational thresholding method and a convolutional neural network for the determination of fiber volume contents of carbon fiber reinforced polyamide 6 from μ CT images
Juliane Blarr | KIT



Highlights

Novel thresholding method and convolutional neural network for fiber volume content determination from 3D μ CT images

Blarr, Juliane, Kunze, Philipp, Kresin, Noah, Liebig, Wilfried V., Inal, Kaan, Weidenmann, Kay A.

- Alternative to experimental determination of FVC through CT image processing
- Convolutional neural network and novel thresholding technique
- Especially for low contrast material combinations
- Average absolute deviations of only about 2.7 % (novel thresholding) and 1.46 % (CNN)

Novel thresholding method and convolutional neural network for fiber volume content determination from 3D μ CT images

Blarr, Juliane^{a,*}, Kunze, Philipp^a, Kresin, Noah^a, Liebig, Wilfried V.^{a,b}, Inal, Kaan^c, Weidenmann, Kay A.^{b,d}

^a*Institute for Applied Materials – Materials Science and Engineering (IAM-WK), Karlsruhe Institute of Technology (KIT), Engelbert-Arnold-Straße 4, 76131 Karlsruhe, Germany*

^b*Fraunhofer Institute for Chemical Technology ICT, Pfinztal, Germany*

^c*University of Waterloo, Department of Mechanical and Mechatronics Engineering, Waterloo, Canada*

^d*Institute of Materials Resource Management, Universität Augsburg, Augsburg, Germany*

Abstract

In order to determine fiber volume contents (FVC) of low contrast CT images of carbon fiber reinforced polyamide 6, a novel thresholding method and a convolutional neural network are implemented with absolute deviations from experimental values of 2.7 % and, respectively, 1.46 % on average. The first method is a sample thickness based adjustment of the Otsu threshold, the so-called "average or above (AOA) thresholding", and the second is a mixed convolutional neural network (CNN) that directly takes 3D scans and the experimentally determined FVC values as input. However, the methods are limited to the specific material combination, process-dependent microstructure and scan quality but could be further developed for different material types.

Keywords: X-ray tomography, carbon fiber reinforced polymers, thermoplastics, low contrast, deep learning

1. Introduction

Owing to their superior lightweight potential through excellent specific strength and stiffness, fiber reinforced polymers (FRP) gain market relevance in engineering materials mostly substituting metals in the automotive, aeronautic, energy or sports sector [1]. While thermosets are still more widespread, there seems to be a trend in the field of matrix systems towards using thermoplastics, since they offer more advantageous recycling properties [2] and rapid processing [3]. Despite being more expensive, the use of carbon fibers instead of glass fibers provides superior mechanical properties concerning most aspects, like tensile and flexural strength as well as Young's modulus [4]. Considering the evaluation of μ CT images, which is one of the most common non-destructive testing (NDT) methods for heterogeneous materials, carbon fibers embedded in polymers prove to be challenging. With their smaller diameter (5 μ m - 7 μ m) and close proximity to polymer matrix materials

regarding density, reaching both sufficient resolution and contrast for good distinction between fiber and matrix is difficult [5, 6].

While continuous FRP, like unidirectional tapes, show the best mechanical properties, discontinuously fiber reinforced polymers allow for high design flexibility at relatively low costs [7]. The latter are in general more challenging to characterize, owing to having a more complex, stochastic microstructure. The complexity of the microstructure of the material in this work arises both from the extrusion part (stochastic fiber length distributions) and the compression molding part of the process (locally differentiating fiber volume contents as well as fiber orientation distributions) (cf. Section 3).

In order to be able to characterize the microstructure quantitatively anyway and to create mechanical material models, necessary quantities for description have developed. Fiber length distributions (FLD) describe how many fibers of different lengths occur in a specimen. The fiber length is relevant to the mechanical properties – only above a certain critical fiber length the maximum stress transfer is possible [8].

The fiber orientation distribution is classically described by means of three-dimensional fiber orientation tensors (FOT) of second or fourth order [9, 10, 11]. FOT

*Corresponding author.

Email address: juliane.blarr@kit.edu (Blarr, Juliane)

URL: https://www.iam.kit.edu/wk/21_2831.php (Blarr, Juliane), https://www.grk2078.kit.edu/members_358.php (Blarr, Juliane)

especially affect the isotropy properties. These tensors can be difficult to determine experimentally because removal of the matrix implies dissolution of the orientation-giving structure. Accordingly, they are determined from μ CT images using structure tensor-based image analysis [12, 13]. Since this evaluation is limited to a certain sample size due to the necessary resolution, there are different approaches to interpolate these local fiber orientation tensors to determine a more global orientation distribution [5, 14, 15, 16].

Probably the simplest sounding quantity at first is that of the fiber volume content (FVC). The percentage of fibers compared to the matrix in a specimen obviously affects the mechanical properties of that specimen. In stochastic microstructures, the fiber content varies to a certain degree throughout a component.

Fiber volume contents can be determined experimentally. Pyrolysis, i.e. ashing of the matrix and subsequent weighing of the fibers in comparison to the total mass, is the most common method (also called thermogravimetric analysis (TGA)). This method works quite reliably for glass fibers [17], but poses difficulties for carbon fibers [18]. The authors tested this procedure for the material used in this paper but encountered problems as well. Either the temperature or the period of exposure to temperature was too high and the carbon fibers were externally attacked. Or one of both or both were too small and then the fibers did not disperse properly because residues of the matrix still held the fibers together.

Instead, there is also the option of chemically dissolving the matrix through acid, which is a more stable method for carbon fibers as well. Neither of these methods, however, is a way to determine FVC for natural fiber reinforced polymers. Natural fibers would be attacked by both processes [19]. Moreover, these methods have other disadvantages: Process-related, fiber volume contents also fluctuate within a sample and exhibit specific characteristics along different axes (compare "shell-core effect" [5, 20, 21, 22]). These cannot be detected by the experimentally destructive determination of the fiber volume content. Several sections would have to be made in order to obtain three-dimensional information. A 3D μ CT image, on the other hand, can also provide information about local volume distributions and thus help to understand process correlations. Furthermore, μ CT scans are often required for fiber orientation analysis anyway, so it would not mean any additional effort to determine fiber volume contents from these scans as well.

Despite all these advantages, there is no standardized computational method known to the authors for reliably determining the fiber volume content of carbon fiber reinforced polymers. A simple threshold to distinguish

between matrix voxels and fiber voxels turns out to be clearly too inaccurate [23, 24, 25], especially for discontinuously carbon fiber reinforced polymers. For this reason, this paper deals with a comparison of methods for the determination of fiber volume contents of carbon fiber reinforced (CF) polyamide 6 (PA6) produced in the long fiber reinforced thermoplastic direct (LFT-D) process [26]. Both thermogravimetric analysis, so ashing the matrix, and chemical dissolution of the matrix in acid followed by volume content determination through the weighting of the fibers were tested experimentally. For the reasons mentioned above, only the results of the chemical dissolution are included in detail in this paper. Furthermore, two computational methods are presented. For this purpose, high-resolution three-dimensional computed tomography images of the CF-PA6 samples are acquired. The novel thresholding method "average or above" (AOA) was developed, which was validated using the experimental FVC results of the chemical dissolution of the matrix. In addition, a convolutional neural network (CNN) was developed to predict the FVC directly from the 3D CT images, which was trained using the scans and also the acid-based FVC values.

2. State of the Art

Gray value thresholding has been an image processing method known for years. Multiple different automated thresholding methods for picture segmentation have been developed; some of the most common are the one by Otsu in 1979 [27], the moment-preserving method by Tsai [28] or the mean threshold [29, 30, 31]. Pre-implemented thresholding methods in ImageJ have been used to determine the fiber volume content of glass fiber reinforced polymers, e.g. in [32]. However, Pinter et al. mention that high contrast between matrix and fiber is required, which is not necessarily given for CFRP. The authors applied conventional thresholding techniques onto the material at hand showing its inapplicability in this case in the methods (Section 4). These common thresholding techniques are still widely used when determining pores, voids or unsaturated and hence dry fiber areas in composites [33]. Air inclusions are usually easy to detect due to their high contrast. Nonetheless, even in the case of entrapped air, deep learning has shown to outperform conventional thresholding techniques, especially concerning the detection of small voids or the detection of voids in volumes with small percentages of porosity [34]. A combination of machine learning (ML) and conventional thresholding has also shown to be an option for low contrast and noisy X ray images [35].

Gandhi et al. introduced a new μ CT procedure as another thresholding based option for FVC determination for glass fiber reinforced polymers in 2020 [36]. While they state that the true threshold value is unknown, they proposed to calculate the threshold value as the histogram midpoint between the mean value representing the fibers and the mean value of the background. Subsequently, they performed a normalization step by dividing the individual fiber volume concentration of each slice by the average fiber concentration of the entire μ CT data set. They claim that this way the ambiguity in selecting the true threshold value is resolved. The authors address the impossibility of application of this technique onto the scans in this work in the methods as well (Section 4).

For the problem of the reliable determination of the FVC of carbon fiber reinforced plastics, deep learning (DL) represents a promising opportunity to improve the status quo. CNNs have shown that they can handle low contrast images and improve statements resulting from those [37, 38, 39, 40, 41, 42]. Necessary filtering and image processing steps do not have to be found and applied by the engineer. In particular, deep learning has been used for the segmentation of low-contrast carbon fiber composites [43].

3. LFT-D process and material

The material used as the basis for the FVC determination in this work is carbon fiber reinforced polyamide 6. It is produced in the so-called "long fiber reinforced thermoplastic direct (LFT-D)" process [26] (process graphic, e.g., in [5]). This process is characterized by its inline compounding of the polymer, direct uncut fiber addition, and subsequent compression molding of the plastificate/initial charge into a plaque/part. Images of two plaques can be seen in Figure 1 and 2. The charge area, where the plastificate was inserted, can be optically differentiated from the flow area that was filled during the compression process.

The LFT-D process has the advantage that it is a fast process compared to similar processes for FRP production and as a thermoplastic matrix is used, the product is recyclable as the matrix can be molten. The fact that the fibers are added to the polymer mass uncut, and are only divided into smaller pieces by the shearing of the screw movement, leads to pronounced fiber lengths distributions and typically slightly higher aspect ratios that in comparable processes. In general, this process causes locally varying characteristics of fiber length, fiber volume content and fiber orientation exacerbating microstructure characterization and homogenization/material modeling.

This is also due to the initial charge already having specific porosity and fiber orientation distributions [44].

For the investigations carried out, specimens from two different plaques are used, that were produced in the same way (hence equal fiber addition, mold temperature, screw speed, etc.). The specimens are cut out of the plaques through waterjet cutting. The two cutting plans are superimposed on an image of a plaque in Figure 1 and 2.

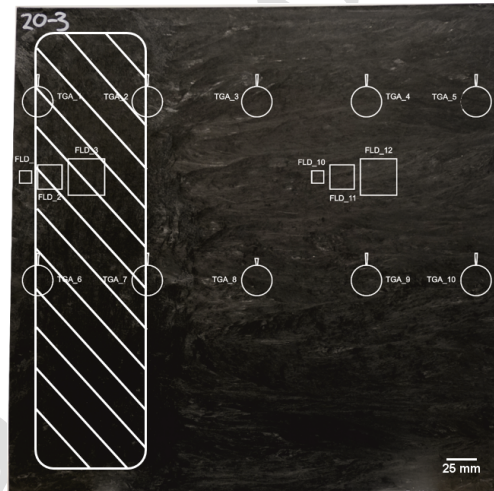


Figure 1: Specimen sizes and positions in the first LFT plaque (cut by waterjet technique) and insertion area of the plastificate at the left (white patterned area). The round specimens were used for pyrolysis investigations and the rectangular specimens for the chemical dissolution of the matrix (cf. the results in Table 3).

It shall be mentioned that the round specimens in Figure 1 were meant for the experimental ashing process, while the rectangular specimens were used for the chemical dissolution of the matrix. The round specimens had a diameter of 25 mm as this was the size of the crucible, in which the TGA was carried out. As an acid-based solution was utilized in the end, the specimens in the second plaque are exclusively rectangular (cf. Figure 2). Additionally, in plaque 1 (cf. Figure 1) multiple sizes of the rectangular specimens are cut: FLD_1 and FLD_{10} have the dimensions 10 mm \times 10 mm \times 3 mm, FLD_2 and FLD_{11} have the dimensions 20 mm \times 20 mm \times 3 mm and FLD_3 and FLD_{12} have the dimensions 30 mm \times 30 mm \times 3 mm. Possible implications of the specimen size on the fiber volume content and fiber length distribution could thereby be detected. As there were no significant changes in FVC between the medium sized sample and the biggest ones, the authors assumed that a

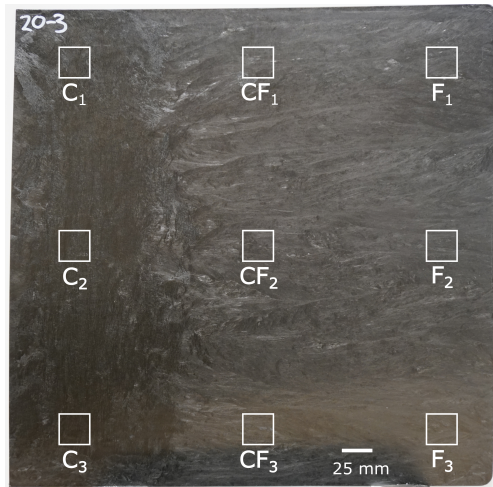


Figure 2: Specimen sizes and positions in the second LFT plaque (cut via waterjet). The results of the FVC measurement after chemical dissolution can be seen in Table 4.

size in between (25 mm side length) should be acceptable as dimension for the specimens of the second plaque. In order to not cut any long fibers, a fiber length distribution measurement was also conducted for the differently sized specimens of plaque 1. The fiber length measurements showed that it is extremely rare for fiber lengths to exceed 10 mm. Furthermore, the on average fiber length increased between the smallest and the medium specimen size but did not in both cases between the medium and the large specimen size. Hence, a saturation of the influence of the specimen size on the fiber length can also be assumed. For the reasons of the FLD measurements, the crucible size of the economically used TGA and the saturation of the FVC adjustment due to the sample size, the side length of 25 mm was chosen for the nine specimens in plaque 2. Additionally, the specimens were extracted from different regions of the plaques. In plaque 1, the three differently sized specimens are cut out of the charge area (FLD_1 to FLD_3), where the initial plastificate is inserted, as well as out of the flow area (FLD_{10} to FLD_{12}). In plaque 2, the specimens from the charge area are labeled with a C, the ones from the flow area with a F and the specimens in-between with CF and a corresponding number, respectively. This way, differences in the microstructure between charge and flow area can also be considered.

The material constituents, with which the final CF-PA6-LFT-D plaques were manufactured, are Zoltek PX35 commercial carbon fibers and a PA6 by the com-

pany DOMO.

4. Methods

4.1. μ CT scans

The Zeiss Metrotom 800 (cone-beam) μ CT system at wbk Institute of Production Science at KIT with the flat panel detector PaxScan2520V with $1536 \text{ pixel} \times 1920 \text{ pixel}$ was used to scan the specimens of the first plaque. The resulting projections were reconstructed to a volumetric image applying the Feldkamp cone-beam algorithm [45] and the Shepp-Logan noise reduction filter. Subsequently, the reconstructed scans were processed in VG Studio Max 3.4.2. The samples of the second plaque were scanned in an YXLON-CT (Yxlon International CT GmbH, Hattingen, Germany) precision μ CT system at Institute for Applied Materials (IAM-WK) at KIT with a μ -focus X-ray transmission tube with tungsten target and a PerkinElmer (Waltham, MA, USA) Y.XRD1620 flat-panel $2048 \text{ pixel} \times 2048 \text{ pixel}$ detector. The scan parameters for the specimens of the two different plaques are listed in Table 1.

Table 1: Scan parameters of the different plaques. Plaque 1 was scanned on the Zeiss device, plaque 2 on the Yxlon device.

Parameter	Unit	Plaque 1	Plaque 2
Voltage	kV	100	110
Current	mA	0.16	0.13
Voxel size	μm	25.98	17.39
Linebinning parameter	-	-	2
Number of projections	-	1450	2220
Exposure/Integration time	ms	1000	800

Exemplary 2D images, so-called slices, of those CT scans are shown in further parts of this work as, e.g., in Figure 10(b).

4.2. Chemical dissolution

The chemical dissolution and subsequent weighting and determination of the FVC was conducted by FIBRE (Bremen). The procedure is briefly described in the following. In order to chemically remove the matrix, approximately 50 ml of concentrated sulfuric acid is added to the samples in an Erlenmeier flask which is then placed on a hotplate (cf. Figure 3).

This is followed by heating until smoke is produced. Subsequently, the samples have to react for one hour at this temperature. After the samples have cooled down ("warm to the touch"), they are mixed with approximately 25 ml of a 35 % hydrogen peroxide



Figure 3: Drop of a specimen in sulfuric acid in Erlenmeier flask and dispersion of the fibers. Courtesy of FIBRE Bremen.

solution and heated until the solution becomes clear and no more gas bubbles rise. The remaining fibers are put in a new specimen cup and can be used for the fiber volume content determination via weighting and for the fiber length distribution determination.

4.3. Computational methods

4.3.1. Application and shortcomings of common techniques

The through-thickness fiber concentration analysis by Gandhi et al. was mentioned in the state of the art [36]. However, the authors were unable to apply this procedure to the μ CT scans in this work as the first step of choosing the midpoint threshold is impossible for the histograms of the CFRP scans as will be elucidated in the next section (Section 4.3.2) and can be seen in Figure 6 (only one peak is visible). The authors did try conventional automatic thresholding methods as comparison to the methods introduced in this work, which can be seen in Tables A5 and A6 in the Appendix (Section 8). Therefore, the two common automatic thresholding methods Otsu (*opencv* [46]) and mean (*scikit image* [47]) were applied, once without filtering the image before and once with the best-performing filter option of our self-implemented method (medianBlur with a kernel size of 15 for plaque 1 and 23 for plaque 2, respectively) applied beforehand. The minimum threshold by *scikit image* did not even compile with a "RuntimeError: Unable to find two maxima in histogram", which confirms the authors' findings. While the two thresholding procedures that worked were applied in Python supporting the subsequent further processing of the values, the exact same threshold and filtering methods can be applied in ImageJ as well. The results were far away from the experimental values. It is noticeable that the results of the pure threshold methods deviate on average by almost 100 % relatively compared to the experimental results. It is particularly striking that the calculated FVC values are almost the same for each sample. A purely constant shift of the threshold value therefore does not appear to

make sense, as this would also not cover the differences between the samples. The use of the median blur at least significantly improved the results with the Otsu threshold. It seems that in the case of low contrast CT images of composites with high fiber volume content, simple thresholding is insufficient for fiber segmentation, which is supported by literature [48].

4.3.2. Novel thresholding method

The novel thresholding method is realized in Python 3.8.7 with the help of the SimpleITK [49, 50, 51], the OpenCV [46] and the NumPy [52] libraries among others. The scans of the samples are generated as 16 bit 3D images in the .raw and .mhd file format. For all further steps those scans were converted into 8 bit. Dark slices at the borders resulting from the specimens not being exactly even-surfaced and further image errors were cut. Each loaded scan was converted into a 3D array. In the following, every slice is handled separately; so the authors iterated through the thickness of the samples and worked on 2D images. At first, a filter was applied to reduce the noise. The filters tested include the "normal" blur filter, the median blur filter, the Gaussian blur filter and the bilateral filter each with various kernel sizes. The kernel size defines the dimension of the window that is slid across the image and in which the filter-specific calculation is performed. The performance of the filters was judged afterwards by comparing the resulting calculated fiber volume contents with the experimental values. The results of the specimens of the first plaque can be seen in Table B7, Table B8 and Table B9 and the results of the specimens of the second plaque in Table C10 and Table C11 in the Appendix (Section 8). The median filter with resolution-adapted kernel size performed the best. It works by creating a kernel of pixels around a central pixel. The values are sorted and the central pixel gets replaced by the median value. From the then noise-reduced image, a threshold value was determined by using the Otsu algorithm. The Otsu algorithm separates an image in two sections by maximizing the inter-class variance

of the gray-level intensities between those sections:

$$\sigma_b^2(t) = \omega_1(t) \cdot \omega_2(t) (\mu_1(t) - \mu_2(t))^2. \quad (1)$$

ω_1 and ω_2 are the probabilities of the two sections with

$$\omega_1(t) = \sum_{i=0}^{t-1} p(i) \quad \text{and} \quad \omega_2(t) = \sum_{i=t}^{L-1} p(i), \quad (2)$$

while $p(i)$ represents the probability for each grey-level intensity. $\mu_1(t) \cdot \omega_1(t)$ represents the mean intensity value of the first section (and vice versa for the second section) with

$$\mu_1(t) = \sum_{i=0}^{t-1} \frac{i \cdot p(i)}{\omega_1(t)} \quad \text{and} \quad \mu_2(t) = \sum_{i=t}^{L-1} \frac{i \cdot p(i)}{\omega_2(t)}. \quad (3)$$

After calculating the Otsu threshold for each slice, it is plotted over the thickness. In Figure 4, the plot for specimen F_1 can be seen as an example.

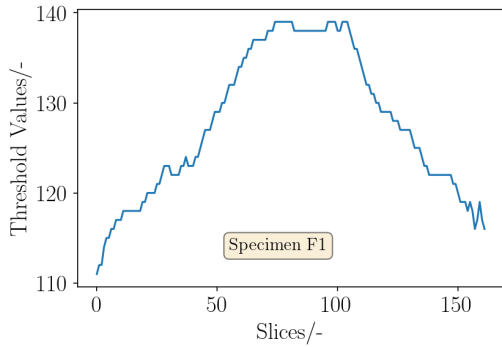


Figure 4: Example of the threshold values from a sample.

Considering the course of the threshold values over the height of the sample, it is noticeable that the threshold values are the highest in the center and drop to a much lower level at the borders of the specimen. Inspecting the scans slice per slice, it is noticeable that there are less fibers visible in those border areas compared to the central layers of the sample as can be seen in Figure 5.

This phenomenon is due to the material flow in the compression molding process. The process-induced difference between outer layers and the center of the part is called shell-core effect in injection molding vocabulary and can be detected considering both fiber volume content as well as fiber orientation [5, 36]. However, this would mean that there are more lighter gray values

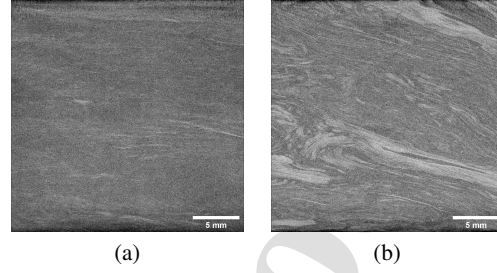


Figure 5: First slice (a) and center slice (b) of the scan of the specimen F_1 . The low amount of fibers and fiber bundles at the border (a) and the contrary high amount in the center (b) is clearly visible.

representing the fibers and less darker ones representing the matrix in the center layers, but the threshold value distinguishing between the two peaks should not shift. This therefore cannot be the main reason for the course of the threshold values over the thickness. However, the authors tested to normalize the threshold values by the average brightness but got little to no improvement on the final FVC results, which seems to support this conclusion.

A second possible explanation would be that of the beam hardening effect, a common phenomenon in computer tomography. The further the beam penetrates the material, the higher the average energy of the photons, as the low energy photons get scattered easily. However, uncorrected images typically show increasing gray values towards the center, hence the rotation axis of the CT. Consequently, this effect would be contrary to the observed one. Additionally, multiple specimens were scanned occasionally, which would superimpose this effect on multiple samples.

Considering the histograms, one understands the issue more. In Figure 6, exemplary histograms of specimen F_1 are given.

On the left (a), the histogram of the entire specimen is shown, in the middle (b), one can see the histogram of one single slice rather towards the border of the specimen and on the right (c), the histogram of a single slice in the center of the specimen is given. In the entire histogram at the left, it is apparent that there are not two peaks as expected. To the contrary, all voxels seem to show gray-values roughly fitting to one single normal distribution. This is due to the bad contrast between carbon fiber and polymer in the CT, which has already been mentioned before (compare e.g. Figure 10 at the right (b)), induced by the closeness of the densities. Additionally, much higher resolutions would be necessary to at least come close to resolving single carbon fibers, which is not given with

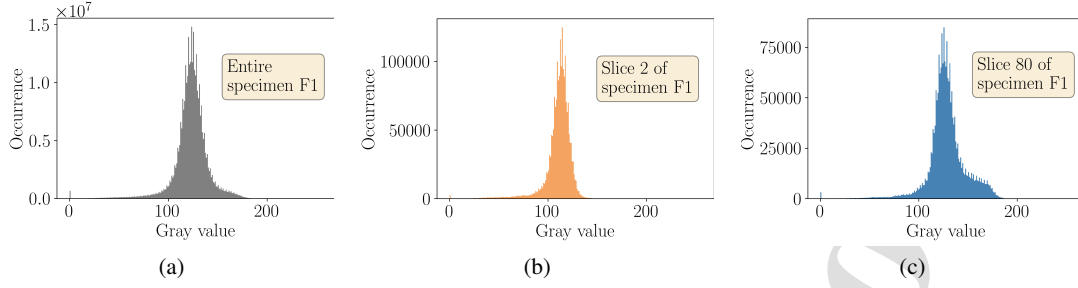


Figure 6: Comparison between the histograms of the entire specimen F_1 (a), a border (b) and a middle slice (c) (after having applied the median filter).

the resolutions of these scans. Both of these facts lead to very noisy images. The Otsu thresholding, respectively any thresholding method for that matter, therefore can not work the way it is supposed to, but calculates some kind of median value of the entire gray value distribution. Looking at the histogram of the border slice, the peak intensity is slightly shifted to the left and there is a small shoulder visible to the left of the distribution. Comparing it with the histogram of the center slice, the distribution is shifted to the right in this case and there is a pronounced shoulder at the right of the distribution. This leads to the rise of the threshold values towards the center of the scans. As these values are more correct than the low threshold values calculated at the borders, there is the need of a non-constant adaption of the threshold values up until the center of the specimens.

Therefore, a two-stage procedure is implemented.

1. The **First stage** consists of the previously described approach. A median blur filter is applied to each slice. Afterwards, the threshold value is determined and saved as $T_{Otsu}[i]$, with i being the corresponding slice.
2. At the beginning of the **Second stage**, the average threshold value

$$T_{average} = \bar{T} = \frac{\sum_{i=0}^{n+1} T_{Otsu}[i]}{n} \quad (4)$$

is calculated. Following, a new array T_{new} is declared. If $T_{Otsu}[i]$ is smaller than $T_{average}$, $T_{new}[i]$ will be set equal to $T_{average}$. Otherwise $T_{new}[i]$ will be set equal to $T_{Otsu}[i]$.

This procedure can be seen in the flowchart in Figure 7.

After that, a binary image is created from each slice i by using the threshold value $T_{new}[i]$. From these binary images, the FVC for each slice can be calculated by determining the percentage of the pixels with a non zero

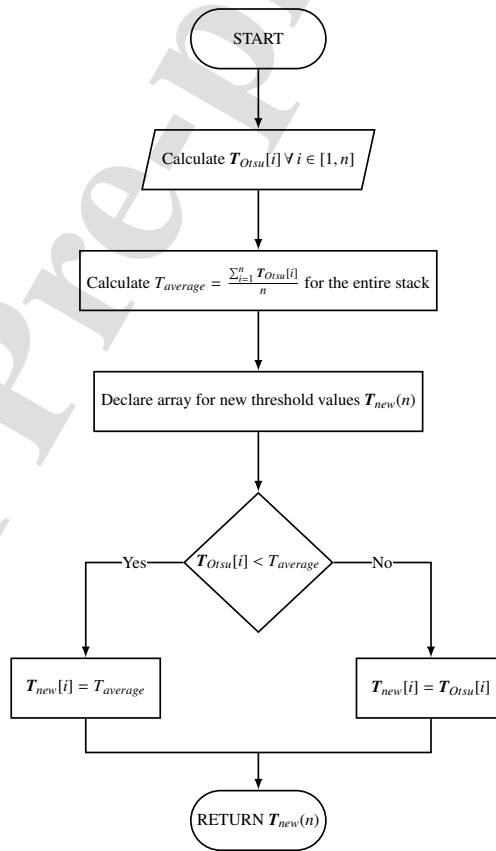


Figure 7: The process of the novel thresholding procedure. At first, the threshold value for each of the n images is calculated with the Otsu algorithm and stored in T_{Otsu} . Afterwards, the average threshold value $T_{average}$ across all slices is determined. Then, the threshold value $T_{Otsu}[i]$ for each image i is compared to $T_{average}$. The larger of the two values is then stored in T_{new} as $T_{new}[i]$.

value. The comparison of T_{Otsu} and T_{new} can be seen in Figure 8.

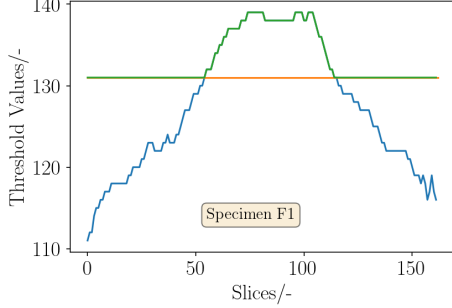


Figure 8: Example of the threshold values from a sample. The blue line represents T_{Otsu} , the orange line $T_{average}$ and the green line represents T_{new} .

The effects of this two-stage approach can be viewed in Figure 9. This empirical procedure was only one among multiple ones tested by the authors, but the one that showed the best results.

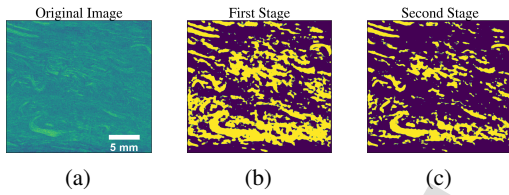


Figure 9: Comparison between the original image (a) and the binary image after the first (b) and second stage (c).

4.3.3. Convolutional Neural Network

Input data and data processing. The convolutional neural network is implemented in Python 3.6.8 with the help of inter alia the NumPy, Scikit-image [47] and SimpleITK packages. Tensorflow [53] and Keras [54] were used as AI framework. The calculations were performed on CPUs provided by the bwHPC cluster 2. For the CNN, the 16 bit scans are loaded directly into the Python script for further processing. In contrast to the thresholding method, the scans are handled as 3D arrays without a loop iterating through the slices. A comprehensive set of uniformization methods is applied so that the neural network's training algorithms solely train the network on the intended differences between the scan data.

For the data loading, uniformization and augmentation, the Python libraries of SimpleITK, Keras and

Numpy provide a large variety of useful methods. However, since the data is processed in 3D, a range of processing methods had to be custom-made. Those helper functions can be found in the Github repository as well. The steps used to process the scans are as follows:

1. Cutting

First, the scans were cropped individually to the actual core material volume to avoid noise at the edges of the scan volume (cf. Figure 10).

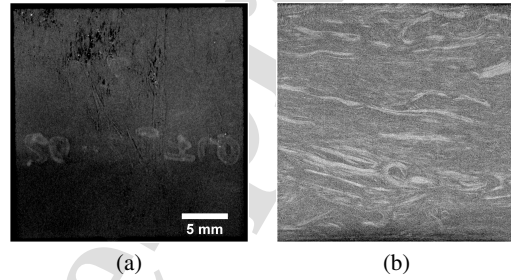


Figure 10: Exemplary slice near the top of the FLD_2 scan (a) showing marking and uneven surface conditions in contrast to (b) a slice of the same scan 10 layers deeper into the material.

The individual amount of cut back slices per specimen can be taken from Table 2.

Table 2: Overview of the original scan data. The resolution is given in terms of the absolute amount of voxels in each dimension. The numbers below "Front/Back cut" refer to the amount of slices removed during data pre-processing.

Scan	FVC	Original scan resolution	Front cut	Back cut
FLD_1	22.3 %	122 x 386 x 386	5	10
FLD_2	25.5 %	128 x 780 x 780	12	12
FLD_3	28.6 %	148 x 1168 x 1162	35	35
FLD_{10}	17.9 %	130 x 391 x 395	9	14
FLD_{11}	24.0 %	135 x 777 x 772	16	16
FLD_{12}	26.6 %	132 x 1164 x 1167	13	15
F_1	23,1 %	168 x 1424 x 1425	0	0
F_2	22,1 %	165 x 1421 x 1425	4	0
F_3	23,1 %	165 x 1416 x 1428	0	0
CF_1	25,6 %	171 x 1403 x 1415	1	5
CF_2	22,3 %	161 x 1422 x 1421	4	0
CF_3	22,8 %	165 x 1406 x 1415	0	0
C_1	26,4 %	165 x 1409 x 1421	2	4
C_2	23,1 %	155 x 1414 x 1421	0	4
C_3	23,8 %	161 x 1406 x 1425	0	4

2. Resizing

All data used to train the CNN should be of the same shape so that one input size of the network can be established. Furthermore, reshaping all arrays into cubes adds an additional possible axis to rotate the data by without changing its shape. That allows for an additional augmentation step and thus doubles the amount of input data after augmentation.

To reshape the cuboid scans into cubes, the *transform()*-method from the *scikit-image* library was used. The target size of the cubes was constrained by the computing power available at the bwHPC cluster. Iterative trials showed that it was capable of executing the script stably up to a cube size of $100 \times 100 \times 100$ voxels, which was thus selected as array dimension.

3. Augmentation

Neural networks require extensive amounts of data to improve their training process. Since only a small amount of scans are available (14), multiple stages of 3D-image augmentations were used to enlarge the input data set.

In the first step, every scan was rotated by 90° and added to the data set with the same FVC as its original. Since the cuboid shape of the transformed data allows for rotation about three independent axes, the process was repeated for the remaining two orthogonal axes. Solely rotating by multiples of 90° ensures that no data is lost at the edges by leaving the scope of the arrays. Furthermore, it is computationally much more efficient than a rotation by a random angle since only the array indexes need to be interchanged.

After multiplying the data set by a factor of 4 by adding rotations of the original scans, all scans are then flipped in a second step. Similarly to the first step, copies of the original scans are mirrored at one plane and then added to the data set with the same FVC as their originals. The process is repeated for the two remaining normal planes, further multiplying the data set by a factor of 8.

Overall, by combining three rotations and three mirroring steps, the amount of input data can be multiplied by a factor of $(1 + n_{rotations}) \cdot 2^{n_{reflections}} = (1 + 3) \cdot 2^3 = 32$. Therefore, the 14 original scans multiply to a data set of 448 samples. More combinations are possible but they lead to exact duplicates of arrays which can be obtained using the method above. I.e. two consecutive 90° rotations about

one axis equal two reflections about two different planes.

4. Split

Before being fed into the neural network, the data, consisting of CT scan arrays coupled with their respective, experimentally observed FVC values, are split into a *training* and a *validation set*. The larger training set is used for the initial tuning of the network's parameters similar to [55].

In between each training epoch, the validation set is used to verify the model's performance on unseen data, which prevents overfitting and allows for an estimation of the model's ability to generalize beyond the training set. More details on the training process are discussed in the *Network training process* paragraph below.

Network architecture. A special characteristic of the network implemented in this work is the direct input of a 3D scan along with a singular scalar value representing its corresponding FVC making it a mixed network. However, the output of the network is only the predicted FVC as a singular value between 0 and 1 for a given scan.

For image processing tasks, the conventional type of neural network is a convolutional neural network (CNN).

The model architecture in this work consists of several layers of convolutions, pooling, dense, and dropout layers. Overall, the neural network can be divided into two stages: The feature extraction stage, where the convolution is happening and the subsequent feature processing stage, where the extracted features are mapped to a corresponding output.

The input layer takes in a 3D tensor of depth, width, and height, representing the CT-scan data, as a single channel since the CT-scans are in gray scale. Furthermore, the single scalar value for the FVC is passed along. The subsequent convolutional layer that extracts features from the input data is followed by a max-pooling layer that downsamples the output of the convolutional layers to reduce the dimensionality of the data and capture the most important features. In this case, a $2 \times 2 \times 2$ max-pooling layer follows the convolution, where out of the $2^3 = 8$ voxels only the largest value is passed on to the next layer. This way, an 87.5 % reduction of data is achieved without a major loss of relevant information since for feature extraction, the precise location of certain features is less relevant. Furthermore, the strongest features are enhanced more as only the highest value inside the kernel area is passed on.

After the convolutional and max-pooling layers, the output is passed through several dense layers with recti-

fied linear unit (ReLU) activation function. In the case of this study, where FVC percentages are evaluated, any negative values are implausible and get filtered out automatically this way. These layers enable the model to learn complex relationships between the input and output. Dropout layers are added after each dense layer to prevent overfitting.

The further one advances into the network during the analysis, the less relevant spacial information becomes, since the ultimate objective is to compute a single scalar, that describes the entirety of a scan. A Global Average Pooling Layer transforms the 3D output of the dense layers into a single-row vector. Its dimensionality corresponds to the amount of feature maps which the feature extraction stage feeds forward. This technique and the choice of Global Average vs. Global Max Pooling was inspired by Zunair et al. [56] and improved the performance of the CNN.

Finally, the output layer is a dense layer with a sigmoid activation function. Its output represents the predicted fiber volume content (FVC) of the CT-scan.

The final network architecture can be seen in Figure 11. The parameters that define the shape of all layers used were determined using parameter sweeps, which are discussed in the *Network optimization* paragraph below. The graphic was made with the help of the latex code published in [57].

Network training process. The CNN described so far is initialized with random weights in all of its layers. Therefore, the initial predictions for the scans of the training set will also be random and are unlikely to show any causal relation to their actual FVC values. In order to tune the weights in a way that enables the network to make reasonable predictions, a training algorithm is used. The training can be divided into two main steps: model compilation and model fitting. During the model compilation step, the configuration parameters are set up. The chosen loss function for this task is binary cross-entropy (BCE). The BCE loss also referred to as log loss or negative log probability is defined as follows

$$l(y, \hat{y}) = -(y \log(\hat{y}) + (1 - y) \log(1 - \hat{y})), \quad (5)$$

with y being the true term (0 or 1) and \hat{y} being the predicted probability (between 0 and 1) [58, 59]. In the perfect case of the model exactly predicting 0 or 1 correctly, the loss amounts to zero. However, in this paper the true term y is not only zero or one, but a continuous value in between. While this is not what the BCE loss was originally intended for, it worked better than other

popular loss functions that are not intended for binary input (and output) such as MSE/MAE or regular cross entropy. Literature also suggests that it works in practice [60, 61], so the authors used the BCE loss in this work. However, it shall be noted that binary cross entropy is asymmetric in the case of the ground truth not being a binary value and the minimum loss is not zero anymore. As an example for $y = 0.24$, which is roughly the average FVC in this case, the minimum loss for $\hat{y} = y = 0.24$ amounts to 0.5511 (cf. Figure 12). This will be relevant when judging the loss plot in the results (Section 5).

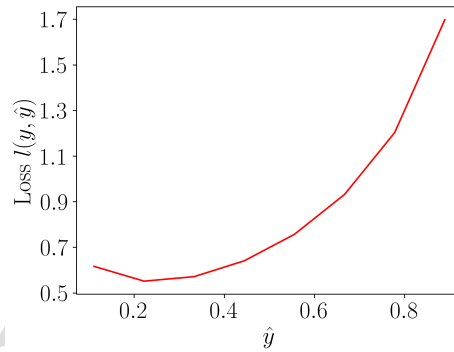


Figure 12: Course of the BCE loss over prediction values \hat{y} between 0 and 1 for a true value of $y = 0.24$.

The optimizer selected for this task is Adam. It is an adaptive learning rate optimization algorithm that adjusts the learning rate dynamically during training, which helps to converge to an optimal solution faster. Because of its computational efficiency and little memory requirement [62], it is a popular choice for training deep neural networks due to its efficiency and effectiveness in updating the parameters of the model. Additionally, the mean squared error (MSE) is used as a metric to evaluate the performance of the model during training. It measures the average squared difference between the predicted and true values. Therefore, it provides insight into the overall accuracy of the predictions of the model along the training process.

After model compilation, the model is fitted to the training data using the *model.fit()* function from Tensorflow. This method implements the general process of machine learning using the parameters set in the compilation process: Using the weights provided at that stage in the network, all training scans are passed through the network. The resulting scalars (the predictions) are subsequently compared to the actual FVC values of the respective specimens. The loss function, which was spec-

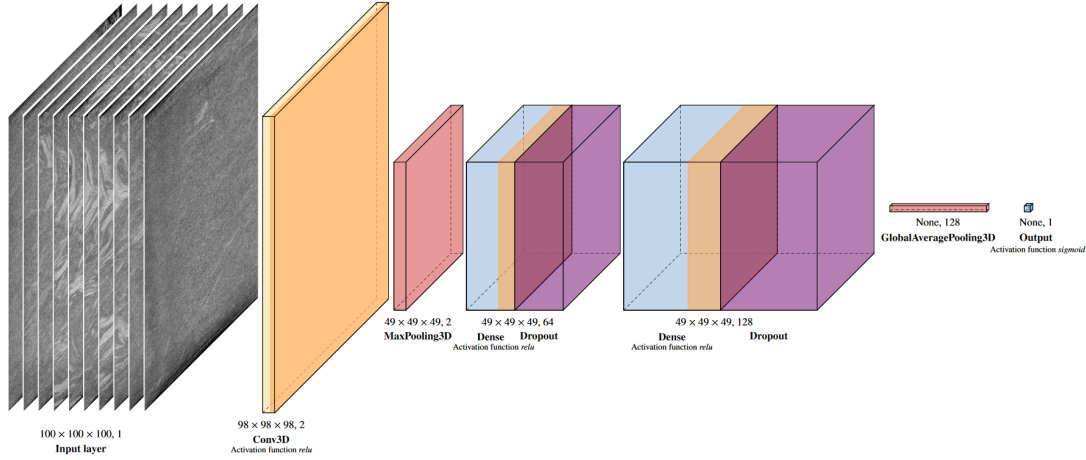


Figure 11: Final architecture of the CNN used in this paper.

ified earlier, takes both the prediction and the real value as its arguments and computes a loss score. The selected optimizer then adjusts the weight in the layers according to the performance of the score, before the entire process is repeated by the set amount of epochs specified in the beginning. Figure 13 describes this process visually.

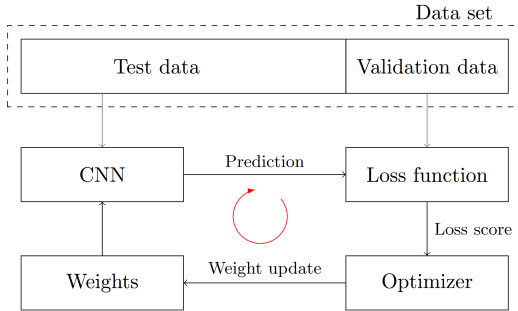


Figure 13: Illustration describing the training of weights in neural networks inspired by [63, 64].

Alongside the training set, the validation set is being evaluated with the same model and loss function simultaneously. Finally, a loss plot such as Figure 17 is generated using the training and validation loss values. The plot shows both as a function of the number of epochs. This visualization helps to monitor the convergence and performance of the model during training, where a decreasing loss indicates that the model is learning and improving over time.

Furthermore, this method also takes over the task of

shuffling the input data and splitting it into training and validation sets, for which a ratio of 2/3 to 1/3 was set. The number of training epochs is set to 40, indicating the number of times the entire training data set is passed through the model during training. Once the model is trained, it is used to predict the fiber volume content for the test data. The deviation between the predicted values and the true labels is then plotted, as seen in Figure 18.

Network optimization. The CNN required a lot of parameters to be defined. The authors started with values provided in similar literature ([55], [64]) and adapted them to the problem at hand. Most parameters were defined by so-called parameter sweeps. Instead of single values, the parameters were provided with a list of values, which were looped through. By changing two parameter values at once, one can find the combination with the best performance. As an example, the amount of filters in a 3D convolutional layer is eligible in Keras. The best amount of connected dense-dropout layers was also unclear. Hence, lists for the amount of filters in the convolutional layer $n_{filters} = \{2, 4, 8, 16, 32\}$ as exponential values to the basis 2 in order to cover a greater field and a list for the amount of dense-dropout layers $n_{dd-layers} = \{1, 2, 3, 4, 5\}$ was given to the network instead of singular values of these parameters. These lists are then iterated through and training is carried out for each combination, i.e. each new network. The standard average deviation for the results of each network is finally compared and the combination of parameters that leads to the best FVC prediction is used.

A similar procedure was followed for the parameters

of the dropout rate in the dropout layer, the pool size in the max-pooling layer and the kernel size in the convolutional layer. Analogously, multiple initial learning rates, loss functions and optimizers were tested until the authors arrived at the presented architecture in Figure 11.

5. Results

5.1. Chemical dissolution

Since the samples that were dissolved in acid of the first plate were of different sizes, the dependence of the fiber volume contents on the sample size can be shown. In fact, this effect seems to be clearly pronounced, as can be seen in Table 3: the fiber volume contents increase monotonically from sample 1 to 3, as well as from 10 to 12.

Table 3: FVC in % for specimens of the first plaque determined through acid-based dissolution of the matrix.

Specimen	FVC
FLD_1	22.3 %
FLD_2	25.5 %
FLD_3	28.6 %
Average	25.5 %
FLD_{10}	17.9 %
FLD_{11}	24.0 %
FLD_{12}	26.6 %
Average	22.9 %

With specimen 1 being the smallest (10 mm × 10 mm × 3 mm), specimen 2 being the second largest (20 mm × 20 mm × 3 mm) and specimen 3 being the largest (30 mm × 30 mm × 3 mm), it is noticeable that the larger the specimen, the larger the FVC. Moreover, the fiber volume content seems to be higher in the charge region (specimen 1, 2, 3) than in the flow region (specimen 10, 11, 12) - a finding that can partly be confirmed considering the samples of the second plate (cf. Table 4). The results of the FVC in Table 4 of the second plaque have been first published by Scheuring et al. [65].

While the average FVC of the three charge specimens (C_1, C_2, C_3) is equal to the global average with 23.57 %, it is indeed slightly higher than the average FVC of the three specimens of the flow region (F_1, F_2, F_3) with 22.74 %. The transition region (specimens CF_1, CF_2, CF_3) shows the biggest average FVC with 24.42 %.

Table 4: FVC in % for specimens of the second plaque determined through acid-based dissolution of the matrix. From Scheuring et al. [65].

Specimen	FVC
F_1	23.07 %
F_2	22.08 %
F_3	23.06 %
Average	22.74 %
C_1	25.57 %
C_2	22.31 %
C_3	22.81 %
Average	23.57 %
CF_1	26.36 %
CF_2	23.10 %
CF_3	23.81 %
Average	24.42 %
Overall average	23.57 %

5.2. Computational determination of the fiber volume content

For the assessment of the overall quality of a method, the following two quantities are introduced. Firstly, the standard average deviation (SAD) between the experimental and predicted FVC values of the validation set (FVC_{exp}, FVC_{pred}) is defined as follows

$$SAD_{abs} = \frac{1}{n_{scans}} \sum |FVC_{exp} - FVC_{pred}|$$

This value describes the **absolute** error of the prediction.

To compare different thresholding techniques with each other, the relative SAD is used at times in this Section given by

$$SAD_{rel} = \frac{1}{n_{scans}} \sum \frac{|FVC_{exp} - FVC_{pred}|}{FVC_{exp}},$$

as a measure of the **relative** error of the prediction.

5.2.1. Novel thresholding procedure

The scans, that the FVC had to be determined of, all showed *salt and pepper* noise. So the first step of the implemented thresholding method was the application of a filter. The choice of the filter type and kernel size was decided on after all other steps of the procedure were defined. The authors conducted a study, as to which filter and kernel size produces a FVC closest to the measured ones. The median blur led to the least average deviation

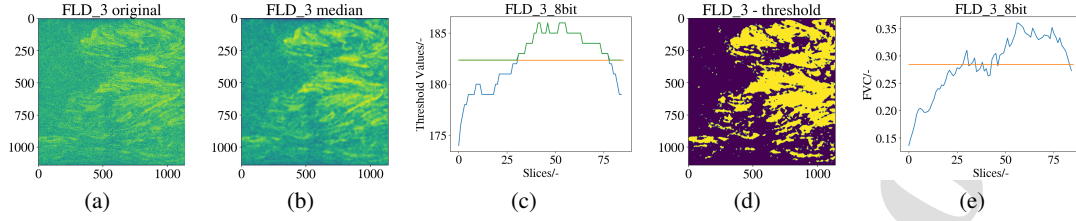


Figure 14: Illustration showing the whole process of FVC determination by AOA thresholding using the example scan of sample FLD_3 . (a) Shown is a middle slice of the original 3D CT image read into Python. (b) The same slice is shown after applying the median filter (kernel size in this case 15, as FLD_3 is one of the first specimens). (c) The Otsu threshold of each slice was determined and plotted over the slices. The calculated thresholds are shown in blue, the average in orange and the value that is ultimately applied in green. The average value is used if the actual threshold of the slice is below the average value, otherwise the calculated value above it is used (cf. Figure 8). (d) The slice is shown with the threshold applied. (e) Finally, the fiber volume content determined by the method is plotted over the slices. The decreasing volume content at the edges is clearly visible. The mean value over all slices in orange is the final determined value, which is again compared with the experimentally determined value.

between the calculated and the measured results. It is noticeable, that there are differences of the best kernel size between the two different plaques. For the specimens of the first plaque, consisting of FLD_1 to FLD_{12} , a median blur filter and a kernel size of 15 pixel \times 15 pixel showed the least average deviation, as well as the least maximum deviation. The results can be seen in Table B7, Table B8 and Table B9 in the Appendix. For the specimens of the second plaque, consisting of C_1 to F_3 , a median blur and a kernel size of 23 pixel \times 23 pixel had the least average deviation, as well as the least maximum deviation, which is shown in Table C10 and Table C11. In the tables with the results of the second plaque, only the tested kernel sizes that seemed the most relevant of some filters are listed. Comparing the dimensions of the images of the different plaques, to be seen in Table 2, it shows that the average image size of the second plaque is larger than that of the first one, hence the image resolution is higher for the second plaque (cf. Table 1). Apparently the needed kernel size of the median blur is dependent on the dimensions and the resolution of the scan used. The effect of the median blur is visible comparing the images (a) and (b) in Figure 14. The determined binary image after applying the AOA thresholding method of this particular slice is shown in (d).

However, not all final results of the FVC are convincing (compare Table B8, column "medianBlur(...,15), and Table C10, column "medianBlur(...,23)). While most samples show decent results, when not including outliers, with a relative average deviation of 1.81 % in the first measurement series and 3.42 % in the second measurement series, there are samples with a much larger deviation. The sample FLD_{10} of the first series and the samples C_2 and F_2 of the second series show relative deviations of up to 116.09 % (FLD_{10}). The reason for those differences of the performances of the algorithm

are not entirely clear. FLD_{10} was a small specimen and a scan with low resolution, which is not the best combination in general. This condition arose because the authors desired to have the same resolution for all samples of one plate and the low resolution was necessary in order to fit the biggest specimens inside of the beam path. Hence, the image quality was insufficient but that was also the case for FLD_1 , which did not deviate that massively. Additionally, FLD_{10} showed a much lower experimentally measured FVC and the authors assumed minor measurement uncertainties leading to that value. For these reasons FLD_{10} was left out of the error calculations of the AOA thresholding and was not used as training data in the CNN either. However, the other two outliers C_2 and F_2 were included. The authors could neither detect any visual deviation nor deviations in the histograms, brightness levels or other measures used for image comparison in these two scans, which is why they are included in the absolute average deviations and were also used as training data for the CNN. That way, the absolute average deviation of the final AOA thresholding with medianBlur filter of 15 and 23, respectively, amounts to about 2.7 %. The final results are depicted in Table D12 in the Appendix.

The final results of the FVC determined by AOA thresholding after applying the median filter as described above are also plotted in Figure 15. The two deviating values of C_2 and F_2 can be clearly detected.

Considering the progress of the fiber content across a specimen by the example of specimen F_2 , a clear non-monotonous course can be seen in Figure 16.

The lower values at the borders of the sample and the increase of the FVC towards the center have been expected. This behavior appears due to the compression molding process. The so-called "shell-core effect" known from injection molding signifies changing fiber

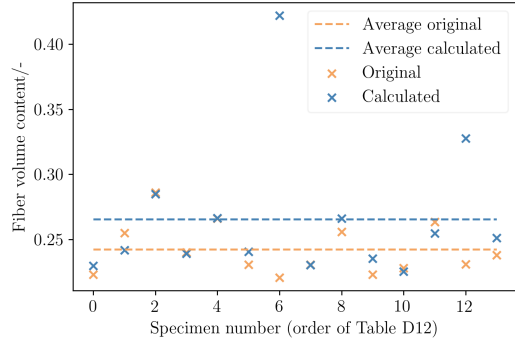


Figure 15: Original (orange) and calculated with the novel thresholding technique (blue) values of FVC for the fourteen specimens, as well as measured averaged FVC (orange dashed) and calculated averaged FVC (blue dashed).

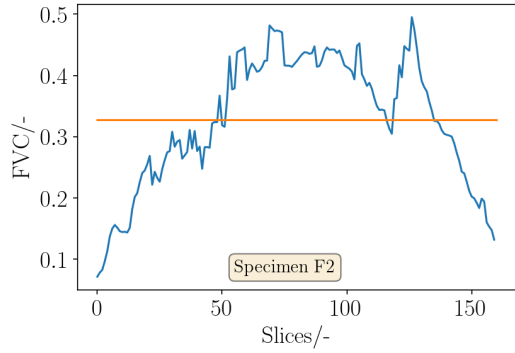


Figure 16: Exemplary progression of FVC over thickness of the specimen considering the example of specimen F_2 . Typical increase of FVC towards the center of the specimen visible.

volume content and fiber orientation between shell and core layer of the plaque [5, 36]. While the fiber orientation effect is only visible in the area of the initial charge (due to a different effect than in injection molding) and non-existent in the flow area, the effect of the process on the through-thickness FVC is clearly visible.

5.2.2. CNN

As shown in Figure 17, using 40 epochs as training of the CNN is sufficient since the model's validation and training losses converge after around 22 epochs and there are no significant improvements beyond that point. The training loss value after 40 epochs amounts to 0.5539 and the validation loss to 0.5535. Considering the shift of the minimum attainable loss value briefly discussed in the methods (Section 4), this can be considered a successful

training process.

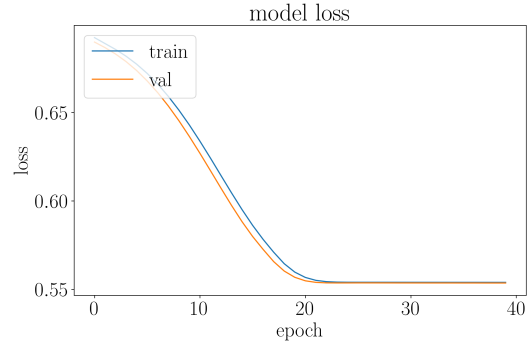


Figure 17: Final graph showing the losses for the testing and validation data sets during the training process as a function of the epoch.

The performance of the CNN is further assessed in Figure 18. Contrarily to Figure 15 which describes the results of the thresholding method, the depicted predicted single values in Figure 18 are themselves an average of the FVC prediction of the CNN for all input scans with the same measured FVC (including the original and the augmented scans).

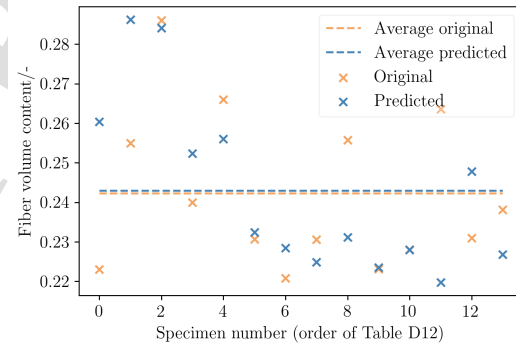


Figure 18: Original (orange) and calculated with the CNN (blue) values of FVC for the fourteen specimens, as well as measured averaged FVC (orange dashed) and calculated averaged FVC (blue dashed).

In the best universal case, hence the final network depicted and described in Section 4.3.3, an absolute average deviation of **1.46 %** was achieved (cf. Table D12).

Beside the amount of epochs used to train the network, a number of variables were tuned to improve the prediction's accuracy. The augmentation process, e.g., has multiple steps of flipping and rotating images. Experiments with using less or more augmentation steps showed that using the most amount of steps showed the

best results. Changing which scans were fed into the program from the start had a large impact and by down-selecting systematically, an absolute average deviation below 0.9 %, so a performance increase by about 40 % was achieved.

However, since the goal of this network is to provide a universal method for carbon-reinforced PA6 sample scans, the authors instead used all data, resulting in worse predictions for the limited validation data at hand in favour of general applicability to different scans at a later time. FLD_{10} , which has already shown to be an outlier for the thresholding method poses an exception in this case. Its implementation caused a remarkable decrease in performance in all cases, for which the authors are lacking an identifiable rationale and thus this scan had to be removed from the entire data set, as mentioned before.

6. Discussion

6.1. Novel thresholding method

With the chemical dissolution average being 24.23 % and the method estimating 26.54 % when including the outliers, the absolute average deviation amounts to 2.7 %, which is higher than that of the CNN. The number was mainly driven up by the two outliers, which were included in the final result as there is no apparent reason to exclude them. Neither the visual inspection of the image nor the histogram showed larger deviations than those that were also usual between the other samples.

The empirical approach of the binary decision between the single slice Otsu threshold and the average stack Otsu threshold could still be developed into another non-constant adaption of the threshold. Specifically, it would probably be helpful to mathematically determine the variation of the histogram across the thickness and change the threshold values accordingly continuously. The rather large influence of the filtering before the AOA thresholding method shall be mentioned here as well (cf. Tables B7, B8, B9, C10, C11). However, the main influence is still the thresholding. Considering Figure 9, it becomes clear that even though a median filter was applied onto both images, the normal Otsu threshold still overinterprets the amount of fibers. The right choice of filter is just an additional factor. Considering in particular that neural networks basically consist of several filters in a row, the importance of choosing the right filter(s) should not be underestimated, especially for noisy, low-contrast images. It shall also be stated that the AOA thresholding method is not optimal by any means. However, the authors wanted to compare the CNN not only

to experimental methods, but also to a more conventional image processing method working on 2D slices and simple thresholding was not an option considering the high errors (cf. Tables A5 and A6). Therefore, multiple approaches were tested, including the mentioned normalization of the threshold values, etc., with the AOA thresholding delivering the best results for our specific material and scan quality. It is a comparably fast and simple method to get to a decent result compared to the high effort in development and training of the CNN. Especially the agile adjustment with the values rising or falling along with increasing or decreasing experimentally determined FVC was convincing.

6.2. CNN

Figure 18 shows the performance of the trained CNN on specific scans of the validation data set. It can be seen that although the FVC of some samples was predicted with very high accuracy, the deviations are distributed unevenly among the data with some predictions being off by several percent points. The overall performance of the CNN is relatively accurate with the average prediction of 24.3 % being very close to the original measured FVC average of 24.23 % and the absolute average deviation amounting to 1.46 % (cf. Table D12). However, it is unreliable in some cases, for which a pattern could not be observed yet. In particular, the logical connection (increasing measured value does not correspond to increasing predicted value from the mean) is not given. However, experimental measurement inaccuracies also fluctuate randomly around a mean value (cf. boxplots).

Although excluding distinct scans from the training data improved the accuracy in this case, it is likely due to the CNN being overtrained on unintended features that are exclusive to the specific data used.

In order to increase the CNN's accuracy for a universal purpose, more diverse input data might be useful. Although a total of 448 scans were used for the training and validation process, they only stem from 14 original scans which themselves contain a similar amount of fibers with FVCs ranging between 22.08 % and 28.6 %. So providing the network with scans that contain significantly more and less fibers could allow it to pick up the relevant patterns more effectively.

Furthermore, a larger scan resolution would detail more information about the structure of distinct fibers, which cannot be detected with the $100 \times 100 \times 100$ pixels. However, any resolution larger than this led to memory issues within the bwHPC cluster and thus posed a limitation on the data evaluation. Rather than only considering bundles of fibers, considering individual fibers would allow the network to differentiate better between fibers

and irrelevant irregularities within the PA6 matrix and could improve the quality of the overall prediction. This would require a CT device with better contrast and higher possible resolutions.

7. Conclusion

The authors implemented a novel thresholding method based on the Otsu threshold and a convolutional neural network both for the computational determination of the fiber volume content of carbon fiber reinforced polyamide 6 produced in the long fiber reinforced thermoplastic direct process. The convolutional neural network outperformed the thresholding in terms of fiber volume content prediction when compared to experimental results that were obtained through acid-based dissolution of the surrounding matrix. The absolute average deviation between the fiber volume content of the average or above thresholding and the experimental values was about 2.7 % compared to about 1.46 % of the convolutional neural network.

8. Outlook

The AOA thresholding method could be adapted to a continuous adjustment of the threshold directly depending on the histogram of the specific slice.

The CNN could be improved by adding additional scans, and hence input data, of which experimental FVC values are available. Especially more specimens with a wider range of different FVCs already in the process would be important as the variance of the experimental data is very much limited. While both methods are highly material-, process-, and scan-dependent, the trained CNN and its determined weights could be used as pretraining weights for a network with a different material type. Furthermore, images of a similar kind, e.g. the same matrix and process but with a glass fiber reinforcement, could be added additionally as input data on top of the CFRP images.

Authors contribution

Juliane Blarr: Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Data curation, Writing - original draft, Writing - review and editing, Visualization, Project administration. **Philipp Kunze:** Methodology, Software, Writing - original draft, Visualization. **Noah Kresin:** Methodology, Software, Writing - original draft, Visualization. **Wilfried**

V. Liebig: Resources, Writing - review and editing, Supervision. **Kaan Inal:** Resources, Writing - review and editing, Supervision. **Kay A. Weidenmann:** Resources, Writing - review and editing, Supervision.

Acknowledgments

The research documented in this manuscript has been funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation), project number 255730231, within the International Research Training Group “Integrated engineering of continuous-discontinuous long fiber reinforced polymer structures“ (GRK 2078). The support by the German Research Foundation (DFG) is gratefully acknowledged. The authors would also like to thank the Fraunhofer ICT for its support in providing the plates produced in the LFT-D process under the project management of Christoph Schelleis and for the TGA investigations by Susanne Lüssenheide. Furthermore, the authors thank the wbk Institute of Production Science and specifically Katja Höger for providing the scans of the first plaque when the institute internal device could not be used. Additionally, the authors would like to thank the FIBRE in Bremen, where the acid-based dissolution investigations were conducted and specifically Ernö Nemeth. Support from Covestro Deutschland AG, as well as Johns Manville Europe GmbH in the form of trial materials is gratefully acknowledged. The authors acknowledge support by the state of Baden-Württemberg through bwHPC.

Data availability

The implemented CNN can be found in the FVC_CNN GitHub repository: https://github.com/jewelsbla/FVC_CNN. The AOA thresholding method can be found in a separate GitHub repository: https://github.com/jewelsbla/AOA_thresholding.

The 3D scan files of all specimens can be accessed and downloaded via the following doi: <https://doi.org/10.35097/1707> [66].

Please cite this paper when using any of the above code or data.

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.

Appendices

A. Results of FVC determined by different conventional thresholding methods

Table A5: Fiber volume contents of the specimens of the first plaque determined by chemical dissolution and by exemplary slice-wise conventional thresholding procedures in Python that are also available in ImageJ. The results of the white row of FLD_{10} were not incorporated into the final maximum deviations, sum of relative deviations and relative average deviations in order to be comparable to the results in Appendix B.

Specimen	Exp. values	opencv Otsu		medianBlur(...,15) + O.		skimage mean		medianBlur(...,15) + m.	
		FVC	FVC	Deviation	FVC	Deviation	FVC	Deviation	FVC
FLD_1	22.30 %	48.90 %	119.28 %	33.18 %	48.79 %	49.58 %	122.33 %	44.10 %	97.76 %
FLD_2	25.50 %	48.89 %	91.73 %	28.20 %	10.59 %	49.66 %	94.75 %	43.50 %	70.59 %
FLD_3	28.60 %	48.93 %	71.08 %	30.89 %	8.01 %	49.69 %	73.74 %	43.55 %	52.27 %
FLD_{10}	17.90 %	49.31 %	175.47 %	44.36 %	147.82 %	49.77 %	178.04 %	47.88 %	167.49 %
FLD_{11}	24 %	48.69 %	102.88 %	32.49 %	35.38 %	49.49 %	106.21 %	43.74 %	82.25 %
FLD_{12}	26.60 %	48.88 %	83.76 %	31.45 %	18.23 %	49.59 %	86.43 %	42.51 %	59.81 %
Max. deviation			119.28 %		48.79 %		122.33 %		97.76 %
Sum of rel. dev.			468.73 %		120.99 %		483.46 %		362.68 %
Rel. aver. dev.			93.75 %		24.20 %		96.69 %		72.54 %

Table A6: Fiber volume contents of the specimens of the second plaque determined by chemical dissolution and by exemplary slice-wise conventional thresholding procedures in Python that are also available in ImageJ.

Specimen	Exp. values	opencv Otsu		medianBlur(...,23) + O.		skimage mean		medianBlur(...,23) + m.	
		FVC	FVC	Deviation	FVC	Deviation	FVC	Deviation	FVC
C_1	23.07 %	47.97 %	107.96 %	30.88 %	33.87 %	49.30 %	113.73 %	44.69 %	93.74 %
C_2	22.08 %	48.78 %	120.91 %	46.33 %	109.82 %	49.47 %	124.04 %	49.00 %	121.91 %
C_3	23.06 %	48.10 %	108.60 %	33.84 %	46.75 %	49.32 %	113.89 %	45.71 %	98.23 %
CF_1	25.57 %	48.32 %	88.94 %	33.25 %	30.01 %	49.37 %	93.05 %	44.16 %	72.68 %
CF_2	22.31 %	47.44 %	112.60 %	36.85 %	65.14 %	49.06 %	119.86 %	45.64 %	104.54 %
CF_3	22.81 %	47.77 %	109.42 %	32.94 %	44.40 %	49.16 %	115.51 %	46.13 %	102.23 %
F_1	26.36 %	48.27 %	83.14 %	39.78 %	50.93 %	49.47 %	87.69 %	46.88 %	77.87 %
F_2	23.10 %	48.98 %	112.01 %	46.58 %	101.62 %	49.71 %	115.17 %	49.00 %	112.10 %
F_3	23.81 %	48.22 %	102.55 %	36.07 %	51.51 %	49.34 %	107.25 %	46.43 %	95.03 %
Max. deviation			120.91 %		109.82 %		124.04 %		121.91 %
Sum of rel. dev.			946.13 %		534.07 %		990.19 %		878.31 %
Rel. aver. dev.			105.13 %		59.34 %		110.02 %		97.59 %

B. Results of FVC of the first plaque determined by AOA thresholding for different pre-filters

Table B7: Fiber volume contents determined by chemical dissolution and by the novel thresholding technique when applying different configurations of the "blur" filter onto the CT images.

Specimen	Exp. values	blur(...,(11,11))		blur(...,(13,13))		blur(...,(15,15))		blur(...,(19,19))	
		FVC	FVC	Deviation	FVC	Deviation	FVC	Deviation	FVC
<i>FLD</i> ₁	22.30 %	25.43 %	14.04 %	23.66 %	6.10 %	22.41 %	0.49 %	21.05 %	-5.61 %
<i>FLD</i> ₂	25.50 %	26.77 %	4.98 %	25 %	-1.96 %	23.76 %	-6.82 %	21.26 %	-16.63 %
<i>FLD</i> ₃	28.60 %	29.83 %	4.30 %	28.95 %	1.22 %	28.36 %	-0.84 %	27.64 %	-3.36 %
<i>FLD</i> ₁₀	17.90 %	36.30 %	102.79 %	35.09 %	96.03 %	34.16 %	90.84 %	32.97 %	84.19 %
<i>FLD</i> ₁₁	24 %	26.07 %	8.63 %	24.65 %	2.71 %	23.26 %	-3.08 %	21.91 %	-8.71 %
<i>FLD</i> ₁₂	26.60 %	26.24 %	-1.35 %	23.58 %	-11.35 %	22.56 %	-15.19 %	20.40 %	-23.31 %
Max. deviation			14.04 %		11.35 %		15.19 %		23.31 %
Sum of rel. dev.			33.30 %		23.34 %		26.43 %		57.61 %
Rel. aver. dev.			6.66 %		4.67 %		5.29 %		11.52 %

Table B8: Fiber volume contents determined by chemical dissolution and by the novel thresholding technique when applying different configurations of the "medianBlur" filter onto the CT images.

Specimen	Exp. values	medianBlur(...,11)		medianBlur(...,13)		medianBlur(...,15)		medianBlur(...,19)	
		FVC	FVC	Deviation	FVC	Deviation	FVC	Deviation	FVC
<i>FLD</i> ₁	22.30 %	26.12 %	17.13 %	24.37 %	9.28 %	22.97 %	3.00 %	21.26 %	-4.66 %
<i>FLD</i> ₂	25.50 %	27.57 %	8.12 %	25.64 %	0.55 %	24.18 %	-5.18 %	22.49 %	-11.80 %
<i>FLD</i> ₃	28.60 %	30.14 %	5.38 %	29.16 %	1.96 %	28.50 %	-0.35 %	27.69 %	-3.18 %
<i>FLD</i> ₁₀	17.90 %	36.30 %	102.79 %	39.47 %	120.50 %	38.68 %	116.09 %	36.81 %	105.64 %
<i>FLD</i> ₁₁	24 %	26.68 %	11.17 %	25.06 %	4.42 %	23.91 %	-0.37 %	22.31 %	-7.04 %
<i>FLD</i> ₁₂	26.60 %	29.20 %	9.77 %	27.56 %	3.61 %	26.64 %	0.15 %	25.79 %	-3.05 %
Max. deviation			17.13 %		9.28 %		5.18 %		11.80 %
Sum of rel. dev.			51.57 %		19.82 %		9.06 %		29.74 %
Rel. aver. dev.			10.31 %		3.96 %		1.81 %		5.95 %

Table B9: Fiber volume contents determined by chemical dissolution and by the novel thresholding technique when applying a "gaussianBlur" and "bilateralFilter" onto the CT images.

Specimen	Exp. values	GaussianBlur(...,(15,15),0)		bilateralFilter(...,15,350,350)	
		FVC	FVC	Deviation	FVC
<i>FLD</i> ₁	22.30 %	28.30 %	26.91 %	24.39 %	9.37 %
<i>FLD</i> ₂	25.50 %	30.96 %	21.41 %	25.56 %	0.24 %
<i>FLD</i> ₃	28.60 %	32.39 %	13.25 %	29.23 %	2.20 %
<i>FLD</i> ₁₀	17.90 %	39.79 %	122.29 %	35.52 %	98.44 %
<i>FLD</i> ₁₁	24 %	28.39 %	18.29 %	25.11 %	4.63 %
<i>FLD</i> ₁₂	26.60 %	29.87 %	12.29 %	25.01 %	-5.98 %
Max. deviation			26.91 %		9.37 %
Sum of rel. dev.			92.15 %		22.41 %
Rel. aver. dev.			18.43 %		4.48 %

C. Results of FVC of the second plaque determined by AOA thresholding for different pre-filters

Table C10: Fiber volume contents determined by chemical dissolution and by the novel thresholding technique when applying different configurations of the "medianBlur" filter onto the CT images.

Specimen	Exp. values	medianBlur(...,15)		medianBlur(...,21)		medianBlur(...,23)		medianBlur(...,25)	
		FVC	FVC	Deviation	FVC	Deviation	FVC	Deviation	FVC
C_1	23.07 %	25.88 %	12.21 %	24.41 %	5.81 %	24.07 %	4.36 %	25.39 %	10.06 %
C_2	22.08 %	41.43 %	87.63 %	42.06 %	90.50 %	42.20 %	91.10 %	42.56 %	92.73 %
C_3	23.06 %	25.48 %	10.52 %	23.43 %	1.62 %	23.04 %	-0.08 %	22.59 %	-2.03 %
CF_1	25.57 %	22.65 %	-11.42 %	25.58 %	0.02 %	26.60 %	4.01 %	28.31 %	10.70 %
CF_2	22.31 %	24.52 %	9.90 %	23.65 %	5.97 %	23.53 %	5.44 %	23.48 %	5.24 %
CF_3	22.81 %	23.33 %	2.28 %	21.97 %	-3.69 %	22.53 %	-1.24 %	22.41 %	-1.76 %
F_1	26.36 %	26.90 %	2.07 %	25.70 %	-2.48 %	25.48 %	-3.33 %	25.28 %	-4.09 %
F_2	23.10 %	32.43 %	40.37 %	32.52 %	40.74 %	32.76 %	41.81 %	35.16 %	52.19 %
F_3	23.81 %	26.46 %	11.15 %	25.29 %	6.23 %	25.11 %	5.49 %	25.00 %	5.01 %
Max. deviation			12.21 %		6.23 %		5.49 %		10.70 %
Sum of rel. dev.			59.55 %		25.82 %		23.95 %		38.89 %
Rel. aver. dev.			8.51 %		3.69 %		3.42 %		5.56 %
Rel. aver. dev. with outliers			20.84 %		17.45 %		17.43 %		20.42 %

Table C11: Fiber volume contents determined by chemical dissolution and by the novel thresholding technique when applying a "gaussianBlur" and "Blur" onto the CT images.

Specimen	Exp. values	GaussianBlur(...,(21,21),0)		blur(...,(15,15))	
		FVC	FVC	Deviation	FVC
C_1	23.07 %	27.03 %	17.20 %	25.86 %	12.13 %
C_2	22.08 %	39.09 %	77.03 %	41.19 %	86.55 %
C_3	23.06 %	26.53 %	15.05 %	24.99 %	8.36 %
CF_1	25.57 %	22.47 %	-12.14 %	21.03 %	-17.78 %
CF_2	22.31 %	25.47 %	14.16 %	24.85 %	11.39 %
CF_3	22.81 %	23.83 %	4.45 %	22.53 %	-1.24 %
F_1	26.36 %	26.00 %	-1.35 %	24.97 %	-5.28 %
F_2	23.10 %	30.93 %	33.88 %	30.41 %	31.64 %
F_3	23.81 %	27.81 %	16.80 %	26.51 %	12.89 %
Max. deviation			17.20 %		17.78 %
Sum of rel. dev.			81.15 %		69.07 %
Rel. aver. dev.			11.59 %		9.87 %
Rel. aver. dev. with outliers			21.34 %		20.81 %

D. Final results of AOA thresholding and CNN (averaged) for the 14 samples and absolute deviations to experimental results

Table D12: Predictions of final AOA thresholding method (compare green columns in Table B8 and in Table C10) and of final CNN structure. The values for the CNN are averaged predictions for the original and all augmented 3D image versions with the same FVC.

Specimen	Exp. values	AOA thresholding		CNN	
		FVC	Absolute deviation	FVC	Absolute deviation
<i>FLD</i> ₁	22.30 %	22.97 %	0.67 %	26.04 %	3.74 %
<i>FLD</i> ₂	25.50 %	24.18 %	1.32 %	28.62 %	3.12 %
<i>FLD</i> ₃	28.60 %	28.5 %	0.1 %	28.41 %	0.19 %
<i>FLD</i> ₁₁	24 %	23.91 %	0.09 %	25.24 %	1.24 %
<i>FLD</i> ₁₂	26.60 %	26.64 %	0.04 %	25.6 %	1 %
<i>C</i> ₁	23.07 %	24.07 %	1 %	23.24 %	0.17 %
<i>C</i> ₂	22.08 %	42.2 %	20.12 %	22.85 %	0.76 %
<i>C</i> ₃	23.06 %	23.04 %	0.02 %	22.49 %	0.57 %
<i>CF</i> ₁	25.57 %	26.6 %	1.03 %	23.12 %	2.46 %
<i>CF</i> ₂	22.31 %	23.53 %	1.22 %	22.35 %	0.04 %
<i>CF</i> ₃	22.81 %	22.53 %	0.28 %	22.80 %	0.01 %
<i>F</i> ₁	26.36 %	25.48 %	0.88 %	21.98 %	4.37 %
<i>F</i> ₂	23.10 %	32.76 %	9.66 %	24.78 %	1.68 %
<i>F</i> ₃	23.81 %	25.11 %	1.3 %	22.68 %	1.12 %
Mean	24.23 %	26.54 %		24.3 %	
Abs. aver. dev.			2.7 %		1.46 %

References

- [1] H. Schüle, P. Eyerer, *Polymer Engineering 2*, Springer Berlin Heidelberg, Berlin, Heidelberg, 2020. doi:10.1007/978-3-662-59841-2.
- [2] H. Schürmann, *Konstruieren mit Faser-Kunststoff-Verbunden: Mit 39 Tabellen*, 2nd Edition, VDI-/Buch, Springer, Berlin, 2007.
- [3] R. Stewart, *Thermoplastic composites — recyclable and fast to process*, *Reinforced Plastics* 55 (3) (2011) 22–28. doi:10.1016/S0034-3617(11)70073-X.
- [4] P. F. Chu, S. Iwasawa, P. L. Schell, C.-Y. Lin, Carbon fiber versus glass fiber reinforcements: A novel, true comparison in thermoplastics, *Polymer Composites* 42 (11) (2021) 6173–6181. doi:10.1002/pc.26294.
- [5] J. Blarr, T. Sabiston, C. Krauß, J. Bauer, W. Liebig, K. Inal, K. Weidenmann, Implementation and comparison of algebraic and machine learning based tensor interpolation methods applied to fiber orientation tensor fields obtained from ct images, *Computational Materials Science* 228 (2023) 112286. doi:https://doi.org/10.1016/j.commatsci.2023.112286. URL https://www.sciencedirect.com/science/article/pii/S092702562300280X
- [6] Y. Sinchuk, O. Shishkina, M. Gueguen, L. Signor, C. Nadot-Martin, H. Trumel, W. van Paepegem, X-ray ct based multi-layer unit cell modeling of carbon fiber-reinforced textile composites: Segmentation, meshing and elastic property homogenization, *Composite Structures* 298 (2022) 116003. doi:10.1016/j.compstruct.2022.116003. URL https://www.sciencedirect.com/science/article/pii/S0263822322007553
- [7] T. Böhlke, F. Henning, A. N. Hrymak, L. Kärger, K. A. Weidenmann, J. T. Wood, *Continuous-discontinuous fiber-reinforced polymers: An integrated engineering approach*, 1st Edition, Hanser Publications, Cincinnati, 2019.
- [8] R. C. Petersen, Discontinuous fiber-reinforced composites above critical length, *Journal of Dental Research* 84 (4) (2005) 365–370.
- [9] S. G. Advani, C. L. Tucker, The use of tensors to describe and predict fiber orientation in short fiber composites, *Journal of Rheology* 31 (8) (1987) 751–784. doi:10.1122/1.549945.
- [10] K. Kanatani, Distribution of directional data and fabric tensors, *International Journal of Engineering Science* 22 (2) (1984) 149–164. doi:10.1016/0020-7225(84)90090-9.
- [11] J. K. Bauer, T. Böhlke, Variety of fiber orientation tensors, *Mathematics and Mechanics of Solids* (2021) 108128652110576doi:10.1177/10812865211057602.
- [12] M. Krause, J. M. Hausherr, B. Burgeth, C. Herrmann, W. Krenkel, Determination of the fibre orientation in composites using the structure tensor and local x-ray transform, *Journal of Materials Science* 45 (4) (2010) 888–896. doi:10.1007/s10853-009-4016-4.
- [13] P. Pinter, S. Dietrich, B. Bertram, L. Kehrler, P. Elsner, K. A. Weidenmann, Comparison and error estimation of 3D fibre orientation analysis of computed tomography image data for fibre reinforced composites, *NDT & E International* 95 (2018) 26–35. doi:10.1016/j.ndteint.2018.01.001.
- [14] C. Krauß, L. Kärger, Tensor interpolation in virtual manufacturing chains for fiber reinforced composites, *International Journal of Mechanical Sciences* 226 (2022) 107378. doi:10.1016/j.ijmecsci.2022.107378.
- [15] J. Blarr, N. Kresin, C. Krauß, K. A. Weidenmann, W. V. Liebig, P. Elsner (Eds.), Application of a Tensor Interpolation Method on the Determination of Fiber Orientation Tensors From Computed Tomography Images, Anastasios, Vassilopoulos, *Composites Meet Sustainability - Proceedings of the 20th European Conference on Composite Materials*, 2022. doi:10.5445/IR/1000155486. URL https://publikationen.bibliothek.kit.edu/1000155486
- [16] J. Blarr, Dataset: FOT interpolation code and scans: 3D μ CT images of specimens of carbon fiber reinforced polyamide 6 plaque, fiber orientation tensor data of these images, and three python code files for two different algebraic and one machine learning based tensor interpolation algorithms (2022). doi:10.5445/IR/1000153725.
- [17] R. Maertens, A. Hees, L. Schöttl, W. Liebig, P. Elsner, K. A. Weidenmann, Fiber shortening during injection molding of glass fiber-reinforced phenolic molding compounds: fiber length measurement method development and validation, *Polymer-Plastics Technology and Materials* 60 (8) (2021) 872–885. doi:10.1080/25740881.2020.1867170.
- [18] T. Abdou, A. Botelho Junior, D. Espinosa, J. Tenório, Recycling of polymeric composites from industrial waste by pyrolysis: Deep evaluation for carbon fibers reuse, *Waste Management* 120 (2021) 1–9. doi:https://doi.org/10.1016/j.wasman.2020.11.010. URL https://www.sciencedirect.com/science/article/pii/S0956053X20306267
- [19] G. Dorez, L. Ferry, R. Sonnier, A. Taguey, J.-M. Lopez-Cuesta, Effect of cellulose, hemicellulose and lignin contents on pyrolysis and combustion of natural fibers, *Journal of Analytical and Applied Pyrolysis* 107 (2014) 323–331. doi:10.1016/j.jaap.2014.03.017.
- [20] C. Schelleis, B. M. Scheuring, W. V. Liebig, A. N. Hrymak, F. Henning, Approaching polycarbonate as an LFT-D material: Processing and mechanical properties, *Polymers* 15 (9) (2023). doi:10.3390/polym15092041.
- [21] M. Bondy, P. Pinter, W. Altenhof, Experimental characterization and modelling of the elastic properties of direct compounded compression molded carbon fibre/polyamide 6 long fibre thermoplastic, *Materials & Design* 122 (2017) 184–196. doi:10.1016/j.matdes.2017.03.010.
- [22] S. Fliegner, *Micromechanical finite element modeling of long fiber reinforced thermoplastics*, Dissertation, Karlsruhe Institute of Technology, Karlsruhe (2015).
- [23] J. S. Rathore, T. Konopczyński, J. Hesser, G. Lucchetta, S. Carmignato, Investigation on tomographic-based nondestructive characterization of short glass fiber-reinforced composites as obtained from micro injection molding, *Journal of Nondestructive Evaluation, Diagnostics and Prognostics of Engineering Systems* 3 (2) (2020). doi:10.1115/1.4046000.
- [24] J. A. Butenegro, M. Bahrami, Y. Swolfs, J. Ivens, M. Á. Martínez, J. Abenojar, Novel thermoplastic composites strengthened with carbon fiber-reinforced epoxy composite waste rods: Development and characterization, *Polymers* 14 (19) (2022). doi:10.3390/polym14193951.
- [25] X.-W. Yu, H. Wang, Z.-W. Wang, Analysis of yarn fiber volume fraction in textile composites using scanning electron microscopy and x-ray micro-computed tomography, *Journal of Reinforced Plastics and Composites* 38 (5) (2019) 199–210. doi:10.1177/0731684418811943.
- [26] W. Krause, F. Henning, S. Tröster, O. Geiger, P. Eyerer, LFT-D — a process technology for large scale production of fiber reinforced thermoplastic components, *Journal of Thermoplastic Composite Materials* 16 (4) (2003) 289–302. doi:10.1177/0892705703016004001.
- [27] N. Otsu, A threshold selection method from gray-level histograms, *IEEE Transactions on Systems, Man, and Cybernetics* 9 (1) (1979) 62–66. doi:10.1109/TSMC.1979.4310076.
- [28] W.-H. Tsai, Moment-preserving thresholding: A new approach,

- Computer Vision, Graphics, and Image Processing 29 (3) (1985) 377–393. doi:[https://doi.org/10.1016/0734-189X\(85\)90133-1](https://doi.org/10.1016/0734-189X(85)90133-1).
URL <https://www.sciencedirect.com/science/article/pii/0734189X85901331>
- [29] C. A. Glasbey, An analysis of histogram-based thresholding algorithms, *CVGIP: Graphical Models and Image Processing* 55 (6) (1993) 532–537.
- [30] T. W. Ridler, S. Calvard, Picture thresholding using an iterative selection method, *IEEE Transactions on Systems, Man, and Cybernetics* 8 (8) (1978) 630–632. doi:10.1109/TSMC.1978.4310039.
- [31] H. J. Russell, Comments on "picture thresholding using an iterative selection method", *IEEE Transactions on Systems, Man, and Cybernetics* 9 (5) (1979) 311. doi:10.1109/TSMC.1979.4310204.
- [32] P. Pinter, Microstructure characterization of continuous-discontinuous fibre reinforced polymers based on volumetric images, Doctoral thesis, Karlsruhe Institute of Technology, Karlsruhe (April 2018).
- [33] I. Tretiak, R. A. Smith, A parametric study of segmentation thresholds for x-ray ct porosity characterisation in composite materials, *Composites Part A: Applied Science and Manufacturing* 123 (2019) 10–24. doi:<https://doi.org/10.1016/j.compositesa.2019.04.029>.
URL <https://www.sciencedirect.com/science/article/pii/S1359835X19301642>
- [34] P. Galvez-Hernandez, K. Gaska, J. Kratz, Phase segmentation of uncured prepreg x-ray ct micrographs, *Composites Part A: Applied Science and Manufacturing* 149 (2021) 106527. doi:<https://doi.org/10.1016/j.compositesa.2021.106527>.
URL <https://www.sciencedirect.com/science/article/pii/S1359835X21002499>
- [35] F. Sosa-Rey, Y. Abderrafai, A. Diouf Lewis, D. Therriault, N. Piccirilli, M. Lévesque, Openfiberseg: Open-source segmentation of individual fibers and porosity in tomographic scans of additively manufactured short fiber reinforced composites, *Composites Science and Technology* 226 (2022) 109497. doi:<https://doi.org/10.1016/j.compscitech.2022.109497>.
URL <https://www.sciencedirect.com/science/article/pii/S0266353822002391>
- [36] U. Gandhi, T. A. Osswald, S. Goris, Y.-Y. Song, *Discontinuous fiber-reinforced composites: Fundamentals and applications*, 1st Edition, Hanser Publications, Cincinnati, 2020.
- [37] J. R. Ferreira, D. A. C. Cardenas, R. A. Moreno, M. D. F. de Sá Rebelo, J. E. Krieger, M. A. Gutierrez, Multi-view ensemble convolutional neural network to improve classification of pneumonia in low contrast chest x-ray images, in: 2020 42nd annual international conference of the IEEE engineering in Medicine & Biology Society (EMBC), IEEE, 2020, pp. 1238–1241.
- [38] Y. S. Moon, B. G. Han, H. S. Yang, H. G. Lee, Low contrast image enhancement using convolutional neural network with simple reflection model, *Advances in Science, Technology and Engineering Systems Journal* 4 (1) (2019) 159–164. doi:10.25046/aj040115.
- [39] R. Kaur, H. GholamHosseini, R. Sinha, Hairlines removal and low contrast enhancement of melanoma skin images using convolutional neural network with aggregation of contextual information, *Biomedical Signal Processing and Control* 76 (2022) 103653. doi:<https://doi.org/10.1016/j.bspc.2022.103653>.
URL <https://www.sciencedirect.com/science/article/pii/S1746809422001756>
- [40] Y. Doi, A. Teramoto, A. Yamada, M. Kobayashi, K. Saito, H. Fujita, Estimating subjective evaluation of low-contrast resolution using convolutional neural networks, *Physical and engineering sciences in medicine* 44 (4) (2021) 1285–1296. doi:10.1007/s13246-021-01062-7.
- [41] L. Zhang, F. Yang, Y. D. Zhang, Y. J. Zhu, Road crack detection using deep convolutional neural network, in: 2016 IEEE international conference on image processing (ICIP), IEEE, 2016, pp. 3708–3712.
- [42] E. Palovcak, D. Asarnow, M. G. Campbell, Z. Yu, Y. Cheng, Enhancing the signal-to-noise ratio and generating contrast for cryo-em images with convolutional neural networks, *IUCr* 7 (Pt 6) (2020) 1142–1150. doi:10.1107/S2052252520013184.
- [43] Y. Sinchuk, P. Kibleur, J. Aelterman, M. N. Boone, W. Van Paeppegem, Variational and deep learning segmentation of very-low-contrast x-ray computed tomography images of carbon/epoxy woven composites, *Materials* 13 (4) (2020). doi:10.3390/ma13040936.
URL <https://www.mdpi.com/1996-1944/13/4/936>
- [44] L. Schreyer, J. Blarr, K. Höger, N. Meyer, L. Kärger (Eds.), *Generation of Initial Fiber Orientation States for Long Fiber Reinforced Thermoplastic Compression Molding Simulation, Composites Meet Sustainability - Proceedings of the 20th European Conference on Composite Materials, Anastasios, Vassilopoulos, 2022*. doi:10.5445/IR/1000155599.
URL <https://publikationen.bibliothek.kit.edu/1000155486>
- [45] L. A. Feldkamp, L. C. Davis, J. W. Kress, Practical cone-beam algorithm, *Journal of the Optical Society of America A* 1 (6) (1984) 612–619.
- [46] G. Bradski, *The OpenCV Library*, Dr. Dobb's Journal of Software Tools (2000).
- [47] S. van der Walt, J. L. Schönberger, J. Nunez-Iglesias, F. Boulogne, J. D. Warner, N. Yager, E. Gouillart, T. Yu, scikit-image: image processing in python, *PeerJ* 2 (2014) e453. doi:10.7717/peerj.453.
- [48] M. J. Emerson, K. M. Jespersen, A. B. Dahl, K. Conradsen, L. P. Mikkelsen, Individual fibre segmentation from 3d x-ray computed tomography for characterising the fibre orientation in unidirectional composite materials, *Composites Part A: Applied Science and Manufacturing* 97 (2017) 83–92. doi:<https://doi.org/10.1016/j.compositesa.2016.12.028>.
URL <https://www.sciencedirect.com/science/article/pii/S1359835X16304560>
- [49] R. Beare, B. Lowekamp, Z. Yaniv, Image segmentation, registration and characterization in r with simpleitk, *Journal of statistical software* 86 (2018). doi:10.18637/jss.v086.i08.
- [50] Z. Yaniv, B. C. Lowekamp, H. J. Johnson, R. Beare, Simpleitk image-analysis notebooks: a collaborative environment for education and reproducible research, *Journal of digital imaging* 31 (3) (2018) 290–303. doi:10.1007/s10278-017-0037-8.
- [51] B. C. Lowekamp, D. T. Chen, L. Ibáñez, D. Blezek, The design of simpleitk, *Frontiers in neuroinformatics* 7 (2013) 45. doi:10.3389/fninf.2013.00045.
- [52] C. R. Harris, K. J. Millman, S. J. van der Walt, R. Gommers, P. Virtanen, D. Cournapeau, E. Wieser, J. Taylor, S. Berg, N. J. Smith, R. Kern, M. Picus, S. Hoyer, M. H. van Kerkwijk, M. Brett, A. Haldane, J. F. del Río, M. Wiebe, P. Peterson, P. Gérard-Marchant, K. Sheppard, T. Reddy, W. Weckesser, H. Abbasi, C. Gohlke, T. E. Oliphant, Array programming with NumPy, *Nature* 585 (7825) (2020) 357–362. doi:10.1038/s41586-020-2649-2.
URL <https://doi.org/10.1038/s41586-020-2649-2>
- [53] M. Abadi, A. Agarwal, P. Barham, E. Brevdo, Z. Chen, C. Citro, G. S. Corrado, A. Davis, J. Dean, M. Devin, S. Ghemawat,

- I. Goodfellow, A. Harp, G. Irving, M. Isard, Y. Jia, R. Jozefowicz, L. Kaiser, M. Kudlur, J. Levenberg, D. Mané, R. Monga, S. Moore, D. Murray, C. Olah, M. Schuster, J. Shlens, B. Steiner, I. Sutskever, K. Talwar, P. Tucker, V. Vanhoucke, V. Vasudevan, F. Viégas, O. Vinyals, P. Warden, M. Wattenberg, M. Wicke, Y. Yu, X. Zheng, TensorFlow: Large-scale machine learning on heterogeneous systems, software available from tensorflow.org (2015).
URL <https://www.tensorflow.org/>
- [54] F. Chollet, et al., Keras (2015).
URL <https://github.com/fchollet/keras>
- [55] J. Pang, W. Sun, J. S. Ren, C. Yang, Q. Yan, Cascade residual learning: A two-stage convolutional neural network for stereo matching, in: Proceedings of the IEEE international conference on computer vision workshops, 2017, pp. 887–895.
- [56] H. Zunair, A. Rahman, N. Mohammed, J. P. Cohen, Uniformizing techniques to process ct scans with 3d cnns for tuberculosis prediction, in: Predictive Intelligence in Medicine: Third International Workshop, PRIME 2020, Held in Conjunction with MICCAI 2020, Lima, Peru, October 8, 2020, Proceedings 3, Springer, 2020, pp. 156–168.
- [57] Haris Iqbal, HarisIqbal88/plotneuralnet v1.0.0 (2018). doi:10.5281/ZENODO.2526396.
- [58] J. Brownlee, Probability for machine learning: Discover how to harness uncertainty with Python, Machine Learning Mastery, 2019.
- [59] M. M. Noel, A. Banerjee, et al., Alternate loss functions can improve the performance of artificial neural networks, arXiv preprint arXiv:2303.09935 (2023).
- [60] F. Chollet, Building autoencoders in keras: The keras blog tutorials (2016).
URL <https://blog.keras.io/building-autoencoders-in-keras.html>
- [61] A. Creswell, K. Arulkumaran, A. A. Bharath, On denoising autoencoders trained to minimise binary cross-entropy, arXiv preprint arXiv:1708.08487 (2017).
- [62] D. P. Kingma, J. Ba, Adam: A method for stochastic optimization (2017). arXiv:1412.6980.
- [63] L. Schöttl, Development of methods for the characterization of damage in sheet molding compounds by using in-situ computed tomography, Doctoral thesis, Karlsruhe Institute of Technology, Karlsruhe (2021). doi:10.5445/IR/1000138851.
- [64] F. Chollet, Deep Learning with Python, Second Edition, Manning, 2021.
URL <https://books.google.de/books?id=XHpKEAAAQBAJ>
- [65] B. M. Scheuring, N. Christ, J. Blarr, W. V. Liebig, J. Hohe, J. Montesano, K. A. Weidenmann, Tensile properties of long fiber-reinforced thermoplastics with mixed fiber types: Experimental investigations and prediction by orientation-based homogenization approaches, Submitted in 2023 to: International Journal of Mechanical Sciences (2024).
- [66] J. Blarr, Computed tomography scans of 15 different carbon fiber reinforced polyamide 6 specimens (both original files (16 bit) and files of further cut scans (8 bit)). Format: .mhd and .raw (2023). doi:10.35097/1707.

Declaration of interests

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.

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

Journal Pre-proof