Single system for online monitoring and inspection of automated fiber placement with object segmentation by artificial neural networks

Marco Brysch[1] · Mohammad Bahar[1] · Hans Christoph Hohensee[1] · Michael Sinapius[1]

Abstract

The reduction of material defects in the automated fiber placement process is one of the significant factors for manufacturing large and complex components more efficiently in the future. However, the monitoring of complex manufacturing processes usually requires complex sensor and computer systems that are often quite sensitive to disturbances and errors. New techniques such as image segmentation with neural networks provide a new approach to this problem and have the potential to solve complex processes faster and more robustly. In this study, a system is presented that performs monitoring, inspection and measurement tasks simultaneously in automated fiber placement processes. The system is based on the SiamMask network which is used for the automatic image processing. The artificial neural network is trained to recognize individual carbon fiber tapes and segment them for additional analysis. For the creation of the testing- and training data, an analytical approach is presented. The post-processing of the object segmentation, which is the primary output of the SiamMask network and the identification of individual tapes, provides accurate measurements which are demonstrated by an example. We show that image segmentation with modern approaches like SiamMask offers great potential to handle highly complex engineering tasks in a faster and more intelligent manner in comparison to conventional methods.

Introduction

Fiber-composite materials are applied in several industrial applications as they offer components with high specific strength and stiffness. Automated techniques such as automated fiber placement (AFP) enable the manufacturing of large and complex parts, especially in the aerospace industry, such as a fuselage that consists of carbon fiber-reinforced polymer (CFRP) (Liu et al. 2016). It has been demonstrated that the AFP process cannot be completely error-free (Rakhshbahar and Sinapius 2018), especially on double-curved surfaces. In this process, gaps, in particular, have a significant effect on material properties. A study by the American Institute of Aeronautics and Astronautics (AIAA) has shown that a gap of 0.76 mm can reduce the compressive

strength of the laminate by up to 27% (Sawicki and Minguet 1998; Heinecke 2019; Ghayour et al. 2020). In this context, the aerospace industry has rigorous tests and standards for the inspection of laminates to check for material gaps and overlaps (Heinecke and Willberg 2019). Continuous quality control of the laminate during and after production is therefore essential.

More than 60% of the total production time is spent on the inspection and reworking especially for large structures (Halbritter and Harper 2012; Meister et al. 2020). In this context, the AFP process in combination with thermoset materials offers a decisive advantage. Faulty or incomplete bonding between the carbon fiber tapes (CFT) can be reworked substantially easier compared to thermoplastic material applications.

By monitoring the AFP process, the position and size of placement effects like gaps can be identified. This information can be used to evaluate laminate quality. A new method allows filling the gaps after its detection simultaneously via 3D printing with continuous carbon fiber (Rakhshbahar and Sinapius 2018). Common gap detection techniques include

✉ Marco Brysch
m.brysch@tu-braunschweig.de

[1] Institute of Mechanics and Adaptronics (IMA), Technische Universität Braunschweig, Langer Kamp 6, 38106 Braunschweig, Germany
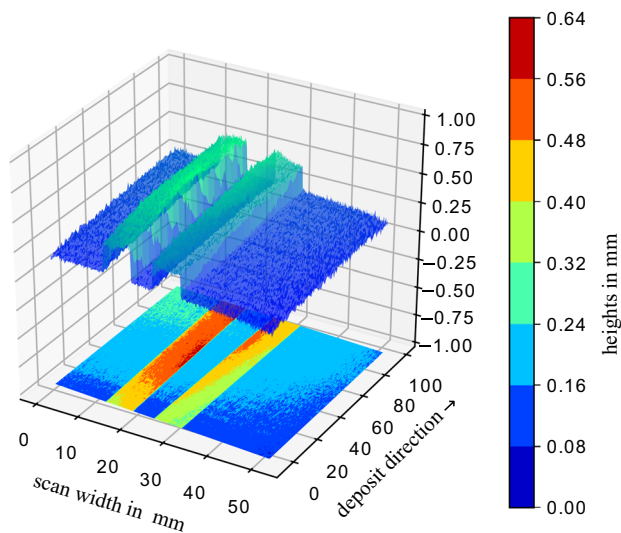
Fig. 1 AFP inspection method for gap detection based on laser profilometer measurements. A three dimensional profile of the CFT's can be created by the recorded profiles in combination with the operation head position

Fig. 2 Heat distribution during the layup process based on a numerical simulation. Individual CFT's, their beginning and end can be identified by the temperature gradient. In practice, the CFT's are determined using thermal imaging

spectral cameras that can detect different bands of the electromagnetic spectrum (Soucy 1996).

This type of inspection is very similar to human vision and depends on the contrast between the faultless and faulty areas. The evaluation of the image data can be improved by lighting systems to improve the contrast conditions (Schmitt et al. 2006). Other methods for gap identification are based on a measuring technique that determines e.g. 3D profile measurements with laser profilometers (Nguyen et al. 2012). The profile sensor can detect the deviation between points to each other or to a reference in Y- and Z-direction. Also, the X-position can be determined based on the position of the profile sensor installed in the robot. Through these data, the position and size of the defect can be simulated. These methods are often used for detecting gaps and overlaps. The data can be used to create an accurate, three-dimensional profile as shown in Fig. 1. However, this requires the exact alignment of the 3D laser profilometer to the object of inspection, which can limit the measurability of complex surface structures (Bahar and Sinapius 2020). The AFP heads are typically controlled by a programmable logic controller (PLC). The coupling between the profile sensor and the PLC as well as the connection with the robot is complex, especially in real-time .

Also, thermographic systems are often used to detect defects in fiber-reinforced components (Schmidt et al. 2017) as shown in Fig. 2. Different methods are used such as optical heat or cold sources, flash- and halogen lamps, and ultrasound or eddy current excitation (Ibarra-Castanedo et al. 2009; Völtzer 2018; Oromiehie et al. 2019). A thermal camera is installed directly after the consolidation roller. T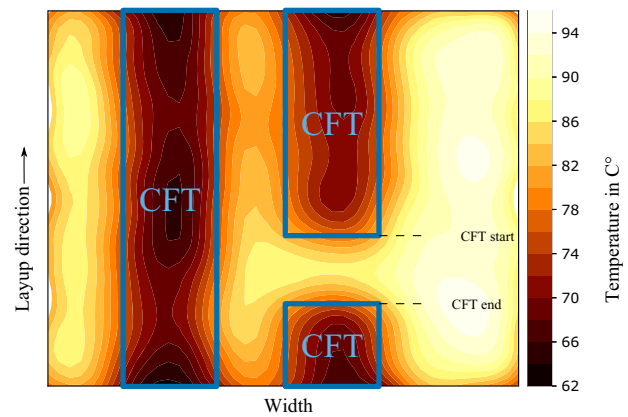he deposited materials are cooled at room temperature depending on the thickness. For example, the cooling rate in gaps is faster than in overlaps. This effect can be used to visualize and detect a defect. However, a thermographic camera can only provide qualitative evidence of various defects, but not quantitative measurements. Optical thermography allows the detection of a wide range of placement effects but is complex and expensive.

The spectral camera can also be used for online monitoring. However, a disadvantage of the spectral camera is that the black prepregs prevent a high visual contrast between a towing tape and the tool surface, which makes it difficult to detect defects (Soucy 1996). The use of multiple light sources can increase the detection of defects. However, this method is limited to the detection of gaps (Tao et al. 2016).

Recent research uses the method of image segmentation based on laser triangulation sensor depth information (Zambal et al. 2019). Image segmentation belongs to the group of computer vision and digital image processing. Classical methods like pixel-thresholding (Sezgin and Sankur 2004; Wu et al. 1982) or Markov random fields (Pieczynski and Tebbache 2000; Plath et al. 2009) have been replaced by newer approaches with artificial neural networks (ANN). In comparison, ANNs achieve better results and higher accuracy in object segmentation. An overview of segmentation methods with ANN is provided by Minaee (Minaee et al. 2021).

Automatic detection of placement effects is currently the subject of research, especially in the field of computer vision. In this context, machine learning and deep learning methods are becoming increasingly important as they work faster and more reliable compared to conventional engineering models. Machine and deep learning methods, therefore, offer great potential for optimization, as visual and manual inspection alone can account for more than 20% of production time (Cemenska et al. 2015). However, the elimination of error

effects in the manufacturing process without interrupting the production process is still an open question. A comprehensive overview of the application of machine learning in composite manufacturing is given in Sacco et al. (2020).

Tendentially more sensors and data are integrated into the AFP depositing process. However, this requires also interacting measuring systems that must be combined into a complex integrated system which in turn may lead to defects in the composite laminates (Sun et al. 2020).

This study shows a different approach, intending to investigate a single monitoring and measurement system based on siamese object tracking (Ondrašovič and Tarábek 2021). Our method uses camera images as primary data input which are segmented and classified. Our approach is based on the SiamMask-Network (Wang et al. 2019; Bertinetto et al. 2016) which, in contrast, to existing methods, creates binary segmentation masks. The segmentation performance of SiamMask allows the calculation of the distance and angle between the carbon fiber tapes and the detection of overlaps, gaps, and their size. Here, the monitoring is primarily utilized for machine control by continuously checking the angle and distance between the CFT's. From the data, automatic real-time machine movements can be derived to optimize the layup quality. The inspection task checks for gaps and overlaps and thus records the distance between the CFT's. Based on this data, decisions can be made whether post-processing is necessary, e.g. by filling the gaps with 3D printing methods to increase the laminate strength. A monitoring and measurement system that collects detailed information about the position and orientation of each CFT also enables the creation of a digital model of the component. The real-time analysis of such digital models permits an improvement of the laminate quality and an automated process that can produce a lighter material (Rakhshbahar and Sinapius 2018; Parmar et al. 2021).

Applied methods

We compared the SiamMask network to an algorithm for edges detection. Both methods are capable to determine the distance between two tapes. In both cases, the edge detection takes place by image analysis. However, while the strictly mathematical approach for edge detection only detects the tape borders, SiamMask is able to provide segmentation masks of both tapes. This segmentation mask can be applied to create a virtual model of the deposited fibers. In the first step, the momentary distance between the fibers can be measured and used to derive machine instructions. This is possible due to the fast processing of the segmentation masks by the SiamMask network. In addition, the large number of recorded masks provides a model of the fiber layer of a component. In the future, a virtual inspection can be carried out on

such a model. Such an inspection provides information about the fiber orientation, possible defects in the component and the determination of the fiber density, since the spacing and orientation of each individual fiber is known.

The SiamMask network

The architecture of the SiamMask network is built upon two parallel networks as shown in Fig. 3. One artificial neural network (ANN) acts as a template- and the other as a detection-branch (Utkin et al. 2021). The initialization of the SiamMask networks requires an initial selection of the object of interest which happens in the first frame of the video. In the present study, this selection step was performed manually. Afterwards, the marked image section is fed to the template branch and serves as a template for processing the remaining images. The output of the template branch corresponds to an abstract representation of the object and is used as a template for searching in the detection branch.

The remaining images are fed individually to the detection branch as search images. This branch uses a similar structure and almost identical weights as the template branch so that a comparable abstraction is performed. By cross-correlating the outputs of both branches all image sections of the search image are compared with the template. The result is an evaluation of the probability where the object of interest can be found in the current image frame. The output of the SiamMask network consists of the final bounding box, a score that indicates whether it is an object or the background of the image, and the final object segmentation.

Automated tape identification with Gaussian-based detection methods

In the task of edge detection in images two general methods may be distinguished. The first method uses the difference of Gaussians (DoG) and searches for local minima or maxima. The second method applies the Laplacian of Gaussian (LOG) and searches for zero crossings (Marr and Hildreth 1980). A DoG approach was chosen, since LOG is known to be sensitive to noise (Basu 2002). Another advantage of the Gauss derivative kernel is that it offers versatile setting options. Since the CFT has a fixed orientation the direction-dependent sensitivity of the DoG for edge detection is a particular advantage. The images and the segmentation masks are used as input to train the SiamMask network.

The presented analytical approach is based on a sliding window method. Instead of detecting edges at once in the entire image $I(x, y)$, one image section after the other is analyzed. Each section $S(x_h, y_w)$ is defined with a height h and a width w. Hence, the pixels of the section $S(x_h, y_w)$ are within the range $0 > x_h > h$ and $0 > y_w > w$. The section slides from left to right with a distance of one pixel until the
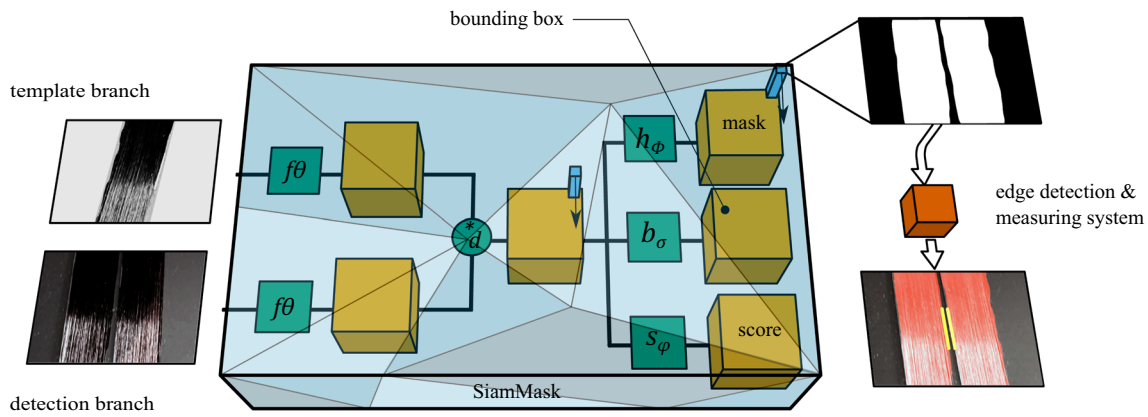
Fig. 3 The SiamMask architecture consists of two *Siamese* networks. Working in parallel, one is acting as template- and the other as detection branch. The segmentation mask generated by the SiamMask network provides a virtual model of the deposited fibers. Based on this mask, the instantaneous distance between the fibers can be measured and used to derive machine instructions. Furthermore, the recorded masks provide a model of the fiber layer of a component on which virtual inspection can be conducted. Such an inspection provides i.e. information on the fiber orientation, possible defects in the component and the determination of the fiber density, since the distance between the tapes is measurable (The image is inspired by Wang et al. (2019))

right end of the image is reached. Afterward, the slide starts again on the left side with a predefined offset. This offset depends on the number of vertical scans given by N.

The overall sectional gradient was computed and taken as a metric M in each section. M is the absolute value of the sum of the sectional derivative computed as

$$M = \sum |g'(x_h, y_w) \circ S(x_h, y_w)|. \qquad (1)$$

The sectional derivative is obtained by the Hadamard product of the Gaussian derivative operator $g'(x, y)$ and the image section $S(x_h, y_w)$ both with size $g', S \in \mathbb{R}^{h \times w}$. The two-dimensional Gauss function, shown in Fig. 4 is given by

$$g(x, y) = A \cdot e^{-\frac{1}{2}\left(\left(\frac{x}{\sigma_x}\right)^2 + \left(\frac{y}{\sigma_y}\right)^2\right)}. \qquad (2)$$

Figure 5 shows the operator $g'(x, y)$ obtained by the total differential of $g(x, y)$, given in Eq. 2, which results in

$$
\begin{aligned}
g'(x, y) &= \frac{\partial g}{\partial x} dx + \frac{\partial g}{\partial y} dy \\
&= -A\left(\frac{x}{\sigma_x^2} dx + \frac{y}{\sigma_y^2} dy\right) e^{-\frac{1}{2}\left(\left(\frac{x}{\sigma_x}\right)^2 + \left(\frac{y}{\sigma_y}\right)^2\right)}.
\end{aligned} \qquad (3)
$$

An advantage of the Gaussian derivative operator $g'(x, y)$ is its versatile filter properties, which can be set through its parameters. Here A is a constant amplification factor and σ_x and σ_y set the sectional blurring level in x- and y-direction. The parameters dx and dy, from the total differential, can be interpreted as sensitivity values, i.e. edges in x-direction will be detected stronger with a higher dx value and vice versa.
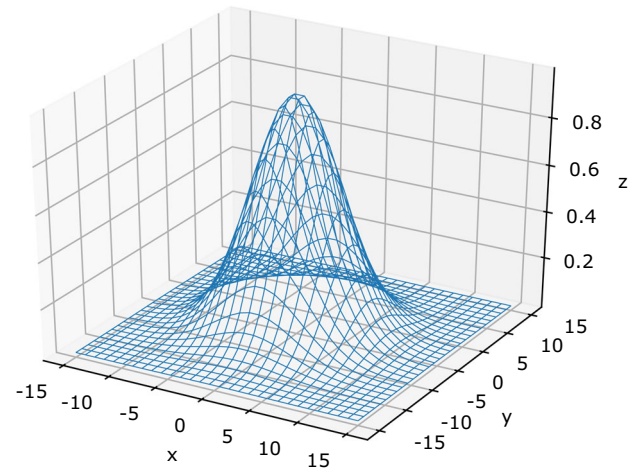

Fig. 4 The two dimensional Gauss function is the calculation basis for the creation of the Gauss derivative kernels

In the present case, the edges were expected to be vertical. Hence a larger dy value is advantageous.

The performance of the segmentation algorithm is tested based on images taken during the layup process. Figure 6 shows such a record during the layup. The images have a size of $w = 1280 \times h = 720$ pixels and a sliding section $w = 30 \times h = 90$. With S as a constant, the variance σ_x, σ_y and dx, dy are defined as the follows

$$\sigma_x = h/S \qquad (4)$$
$$\sigma_y = w/S \qquad (5)$$
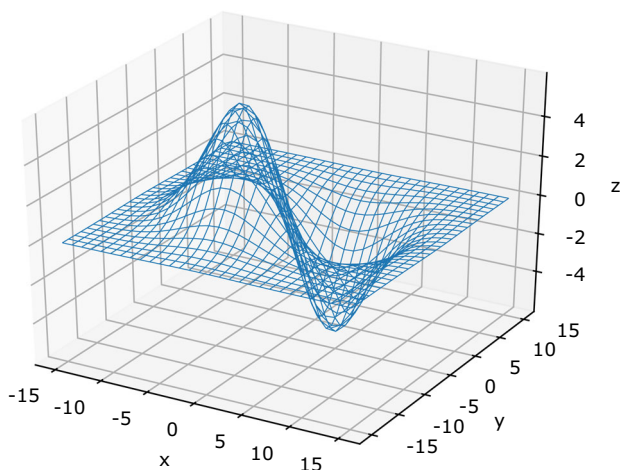$$dx = h \qquad (6)$$
$$dy = w, \qquad (7)$$

Fig. 5 Total derivative of the two dimensional Gauss function given in Eq. 3. This kernel offers versatile filter settings and reliable edge detection
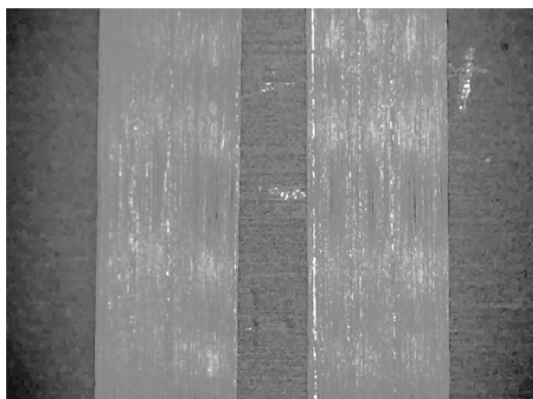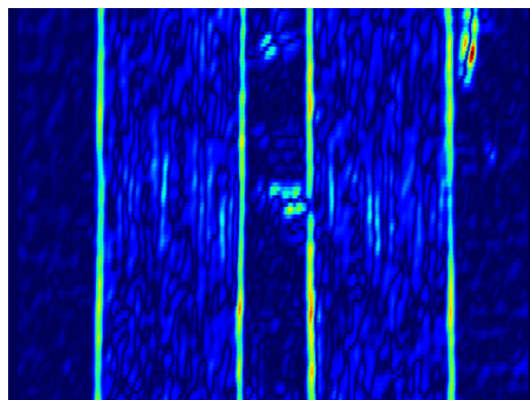

Fig. 7 The sliding window method in combination with the Metric M provides a clear basis for evaluating the pixel gradients. However, it can also be seen that small disturbances quickly appear as gradient peaks, see top right

Fig. 6 The initial images in the grayscale color range had a size of $w = 1280 \times h = 720$ pixels and were recorded with a macro camera during the prototype machine operation

leading to the final form of the Gaussian derivative operator:

$$g'(x, y) = -AS^2 \left(\frac{x}{h} + \frac{y}{w} \right) e^{-\frac{S}{2} \left(\left(\frac{x}{h} \right)^2 + \left(\frac{y}{w} \right)^2 \right)}. \tag{8}$$

For the image analysis, the two constants A = 1 and S = 7 have been defined. An interval x and y was applied with

$$[-h/2, h/2] = \{x \in \mathbb{R}^h | -h/2 \le x \le h/2\} \tag{9}$$

and

$$[-w/2, w/2] = \{y \in \mathbb{R}^w | -w/2 \le y \le w/2\} \tag{10}$$

respectively. For each taken image, the edge detection metric M was computed as shown in Eq. 1. A heat-map from the image was created by collecting all M which can be seen in Fig. 7 which depicts the strength of the sectional gradient.

In general, a higher gradient can be interpreted as an edge. However, the example of the local peak shown in Fig. 7 (right top) also demonstrates that a clear edge identification can be easily polluted by other objects. To create an accurate segmentation mask for the CFT these data must be filtered out.

The data is filtered by selecting the peaks for each data row n in the x-direction, see Fig. 8 (single data row), of the heat-map in which $0 < n < N$. For the selecting process, a threshold was chosen which selects only the highest 5% of the data peaks. The selected peaks are shown in Fig. 8 (selected peaks), which are then separated during the peak sampling process. Also, linear regression is applied to all peaks to obtain a line that represents the tape edges. Finally, a segmentation mask can be obtained by filling the area between two adjacent lines. This is done by connecting the edges to a polygon which represents the final segmentation for each CFT. The average processing time to create a segmentation mask for each image takes about 5.5 s.

Error cases during tape deposition

Apart from determining the tape distance, the segmentation mask also allows to determine the deposit position. For each pixel line, from left to right, it is possible to identify the CFT edge with a rising- or falling edge detection. Three cases can be distinguished by the number of determined edges and their distance, as shown in Fig. 9:

Case 1 **Overlap**: Number of detected edges = 2, if the distance between the two edges is smaller than 2 times the tape width there is an overlap, see Profile 1 in Fig. 9.
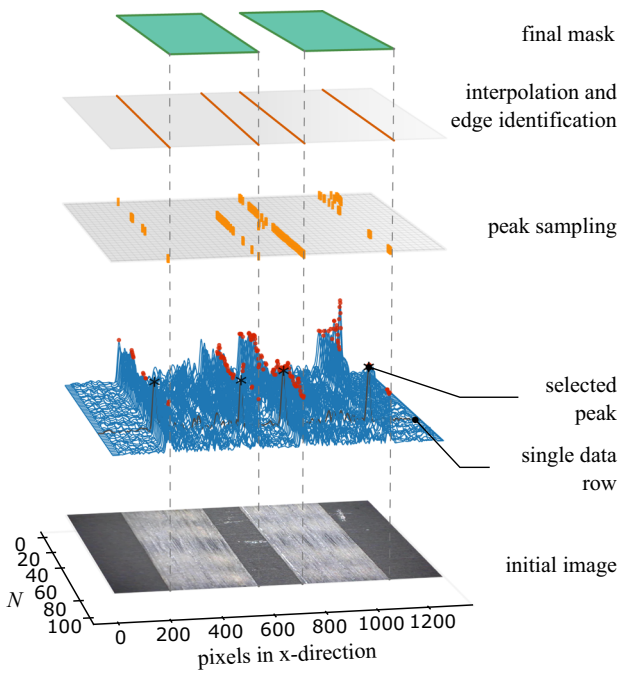
Fig. 8 A Gauss derivation is performed in combination with a sliding-window method from the initial image. The peaks of the sum of the slide derivatives show strong gradients between the slides. These are usually tape edges. Through these peaks, a segmentation mask is created by a linear approximation that covers and identifies the CFT. In this way, outliers are filtered out. The final step consists of filling the area between two adjacent lines to obtain the final fiber segmentation

Case 2 **Fit**: Number of detected edges = 2, if the distance between the two edges is equal 2 times the tape width there is a good fit, see Profile 2 in Fig. 9.

Case 3 **Gap**: Number of detected edges = 4, there is a gap between the tapes, see Profile 1 in Fig. 9. In this case, the size of the gap must be determined.

To reduce the error-proneness three pixel lines lpx_1, lpx_2, lpx_3 were selected from each segmentation mask as shown in Fig. 10. For each line, the edges of the CFT were computed with a rising-/falling edge detector. This gives three points $p_1...p_6$ on the inner edge of each tape . Since one can assume that the CFT has a straight edge in a small area, the three points of an edge are interpolated linearly. This results in two lines l_1, l_2 whose angle and distance can be determined. In the present experiment, the center of the image is selected as the measuring axis. Figure 11 shows the real-time measurement based on the segmentation masks and the two approximated lines in the inner side of the CFT's.

Results

The SiamMask network is trained using a series of images taken during the laying process. The performance and accu-
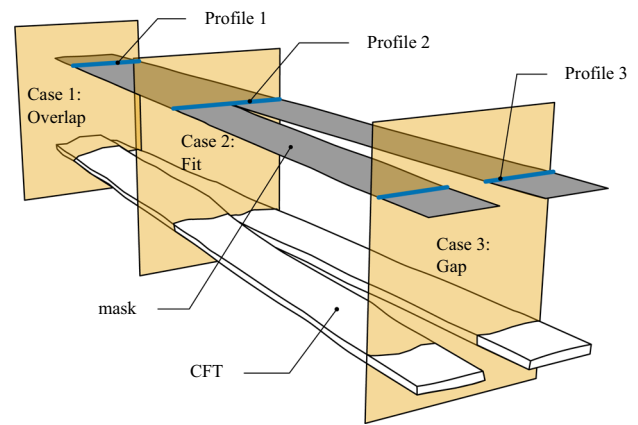
Fig. 9 In the ideal case, the CFT's lie next to each other without gaps or overlaps. However, these errors usually occur during the layup process. With the masks calculated by SiamMask, it is possible to conclude these 3 cases (Profile 1: overlap, Profile 2: ideal fit, and Profile 3: gap). This is done by determining the number and distance of CFT edges
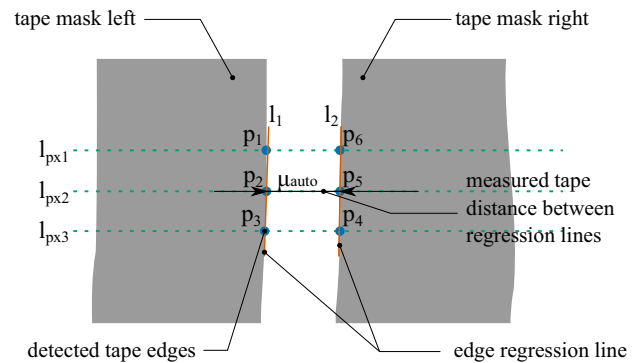
Fig. 10 For each CFT mask, three points were determined on the inside of the edge $p_1...p_6$. Through these points, a linear approximation was calculated. This yields to line l_1 for the left tape and line l_2 for the right tape. The center distance of these lines gives the distance of the CFT's in the middle of the image. This approximation reduced the error caused by the scaling of the segmentation mask

racy of the tape detection and automatic measurement are then compared to manual measurements.

Training data preparation

For the training of the SiamMask network, a large set of training data is required. This data consists on one hand of images that has to be analyzed and on the other hand of segmentation masks of the CFT that need to be recognized by the network. The manual preparation of the segmentation masks was not feasible because of the large dataset. For this purpose, the Gaussian-Based detection algorithm of Sect. 2.2 and the following method for generating masks were applied. This made it possible to create a large number of masks from image captures during the depositing process in a semi-automatic way.
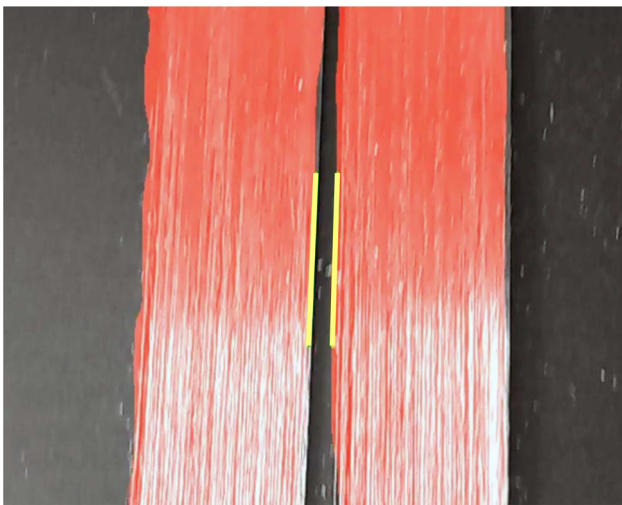
Fig. 11 Snapshot of the segmentation mask (red) and video image with a real-time approximation of the CFT's inner edges (yellow lines) and simultaneous measurement

The automatically generated masks were checked manually as a final step.

Besides a large datasets ANN's also need a variety of data with different case scenarios (patterns) to generalize the problem, i.e. to learn from the data. For this purpose, a matrix is defined with different CFT layup variants that occur during normal machine operation as shown in Table 1. To create the database, 100 video sequences were recorded. These sequences show a layup process of a fixed length. Here, the offset d_o of the CFT's is categorized to the center of the image and the angle α to the image edge. In each case, the gap d_s at the start and the gap d_e at the end of the layup process were measured. Each video sequence consists of approximately 80 single images. In total, 8065 images are taken during machine operation which can be classified into the defined patterns.

Network training and automated tape identification

The training took place at the phoenix cluster of the Technische Universität Braunschweig, equipped with 8x GPU nodes, each with four *NVIDIA Tesla P100 16GB HBM2* GPU's. SiamMask was trained with the ResNet50 (He et al. 2016) backbone. To accelerate the training process, pretrained weights of the visual object tracking database (VOT) (Kristan et al. 2016) were applied. The pre-trained backbone enabled an average training time of less than 16 min for the complete CFT-dataset. The CFT-dataset contains 8065 pictures with corresponding segmentation masks. As described in Sect. 2.2, the images and segmentation masks were processed in 100 sequences in which the sequences were divided into 20 patterns given in Table 1.
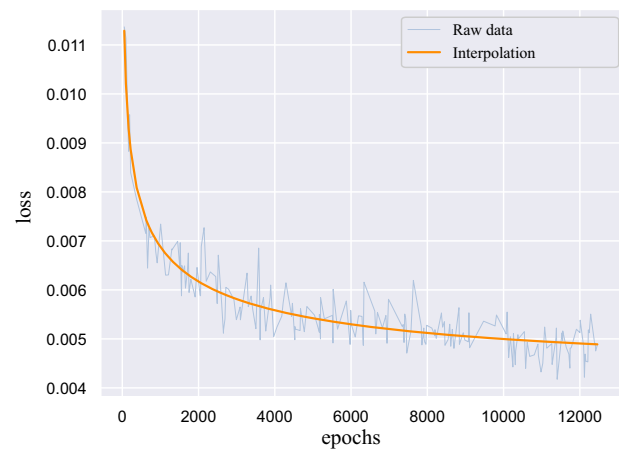
Fig. 12 The loss L_{3B} shown in Eq. 11 converges after 20 epochs to about 0.028
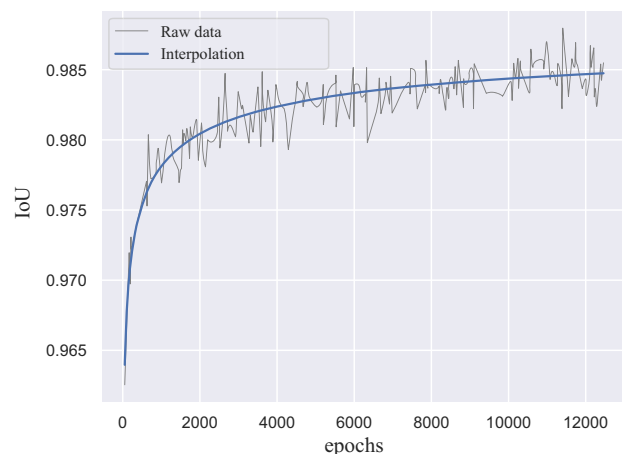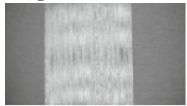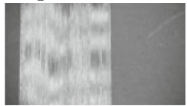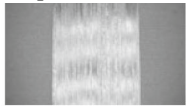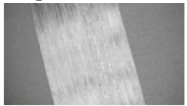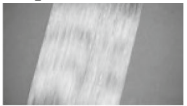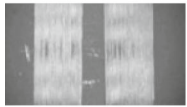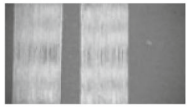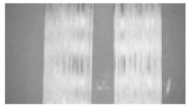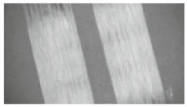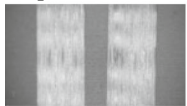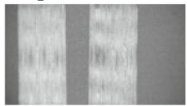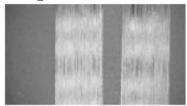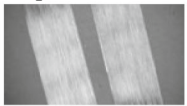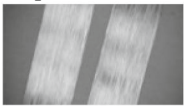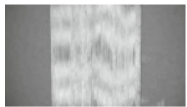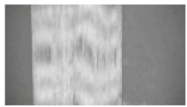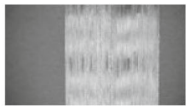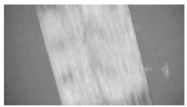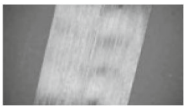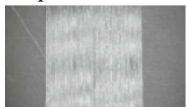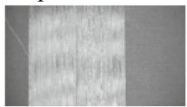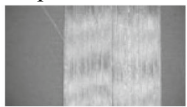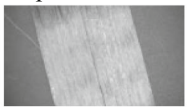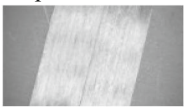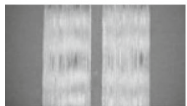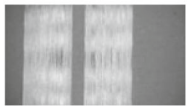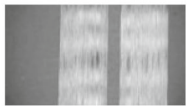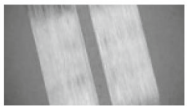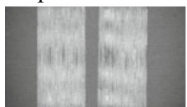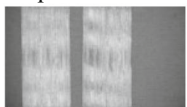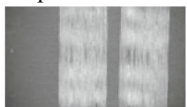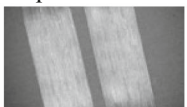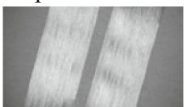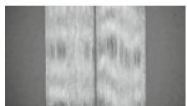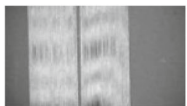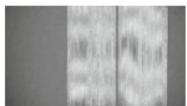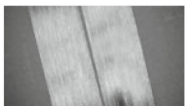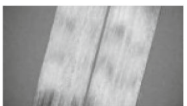
Fig. 13 Training curve for the mean intersection over union (IoU) over 20 epochs shows an accuracy convergence of approximately 98% for the CFT-dataset

The CFT-dataset is split in 70% training- and 30% validation-data respectively. Here, the 20 sequence patterns for the training and validation data were randomly distributed. Data augmentation was not required since the training process achieved sufficient accuracy. After a training time of 20 epochs, the loss L_{3B} given in Eq. 11 converged to about 0.028, as shown in Fig. 12.

Each epoch contains 125 iteration steps. The accuracy shown in Fig 13 reached after convergence approximately 98%. It was measured with the mean intersection over union (IoU), also known as the Jaccard index. This is a measure of the correspondence between the segmentation mask created from SiamMask and the automatically created segmentation obtained with the Gauss derivation method shown in Sect. 2.2.

The Loss and metric parameters were adopted from Wang et al. (2019). Here, a logistic-loss was applied in the calculation of the segmentation mask-loss L_{mask}. The optimization

Table 1 The training data is divided into 20 sequences each representing a variant of the CFT depositing process

	$d_o = 0$mm $\alpha = 0°$	$d_o = -2$mm $\alpha = 0°$	$d_o = +2$mm $\alpha = 0°$	$d_o = 0$mm $\alpha = -15°$	$d_o = 0$mm $\alpha = 15°$
$d_s = -2$mm / $d_s = +3$mm	Sequence 1	Sequence 2	Sequence 3	Sequence 4	Sequence 5
$d_s = +3$mm / $d_s = -2$mm	Sequence 6	Sequence 7	Sequence 8	Sequence 9	Sequence 10
$d_s = 0$mm / $d_s = +1$mm	Sequence 11	Sequence 12	Sequence 13	Sequence 14	Sequence 15
$d_s = +1$mm / $d_s = 0$mm	Sequence 16	Sequence 17	Sequence 18	Sequence 19	Sequence 20

Here the position d_o and the angle α of the CFT were varied. For each of the 20 sequences, 5 recordings of the same type were made. The training data set thus consists of 8065 images in total, each with its individual segmentation mask

of this loss function corresponds to logistic regression. A cross-entropy was applied in the loss definition of the bounding box L_{box} and the score L_{score}. The loss function is defined as

$$L = \lambda_1 \cdot L_{mask} + \lambda_2 \cdot L_{score} + \lambda_3 \cdot L_{box} \qquad (11)$$

with $\lambda_1 = 31$ and $\lambda_2 + \lambda_3 = 1$. The adaptive learning rate was set to a start value of $l_{rs} = 0.01$ and an end value of $l_{re} = 0.0025$.

Real-time segmentation abilities

For the evaluation of the training success, the segmentation results of a trained and an untrained network are compared.

As shown in Table 2, the trained network achieved an IoU (Jaccard index) of 98.4% regarding to Fig. 13. In comparison, an untrained network reached only 87.1% match between segmentation- and label-mask.

The effect of the training becomes clear when looking at the segmentation masks shown in Table 3. In the majority of cases, the untrained network perceives both CFT as one object. Hence, in the untrained state, the network is not able to detect CFT's reliably.

In contrast, the trained network generates relatively accurate segmentation masks. The CFT's are identified as individual objects, gaps can be determined accurately and an overall better segmentation is achieved.

The comparison between the input image and the created object segmentation (in red) is shown in Fig. 14. For the

Table 2 The network IoU increased around 9% after training and it finally achieved a general correspondence of object segmentation and label of over 96%

Weights	Frames per second (fps)	Intersection over Union (IoU)
Untrained	8.03	0.871
Trained	7.99	0.964

Table 3 The comparison between the training segmentation masks (left column), the segmentation with an untrained network (middle column), and a trained network (right column) shows that the training has achieved a significant quality gain

Fig. 14 The evaluation of the object segmentation with a trained network allows identification of the CFT and measurement of the distance in between. Furthermore, multiple gap measurements in one segmentation mask allow an accurate angle determination

following experiment, the distance of the CFT's in the middle of the image is calculated. This is done by counting the pixels between the CFT's and calculating their distance in mm. In the present study, one pixel has a size of approximately $m_p = 0.0172$ mm, which is also the smallest measurable quantity.

A closer look at the created segmentation masks in Table 3 and Figs. 15 and 14 shows that all segmentation masks have irregular edges. These are expected to have a great influence on the measuring accuracy. The irregularities are caused by a scaling process as the network generates the mask as images with a resolution of 127×127 pixels. The segmentation masks are then scaled to the higher resolution of the initial image (in the present case to an image of 1280×720 pixels).

Optimization of the scaling process could improve measurement results for methods with higher accuracy requirements. This can be done either by redrawing the segmentation mask, e.g. with polygons, or by modifying the backbone (ResNet50) to achieve higher object segmentation resolution. Approaches for a higher resolution representation in Siamese meshes can also be achieved with feature fusion on multiple layers, which can achieve a better object localization (Zhang et al. 2020; Zhu et al. 2021).

For a field experiment, the processing rates and the real-time capability between the network and analytical approach were compared. To test the practical use of the network, the

excellent performance of a computer cluster was deliberately omitted. The study was conducted with a standard laptop without GPU support. The implemented hardware consists of an *i7 Intel* of the 8th generation. The results showed that the network took in average 0.9 second per frame including the gap measurement. An analysis of the runtime profile of SiamMask reveals that the program spends about 60% of the time on convolution operations. This time can be significantly shortened with a GPU or TPU. In contrast, the analytical approach without gap measurement requires approximately 5.5 second per frame, see Sect. 2.2. This shows that neural networks offer a feasible method for intelligent and automatic

Fig. 15 The rescaling process of the segmentation masks causes irregular (image top) or blurred (image top) edges which has a great impact on the determination of the gap dimension. This occurs since the measurement accuracy is based on the number of pixels between the CFT edges. An anti-aliasing technique of the CFT edges could increase the accuracy of the inspection system. However, this approach goes beyond the scope of the present study. For comparison, one pixel has a size of approximately $m_p = 0.0172$ mm which is also the smallest measurable quantity

data processing. Moreover, in the present case, the SiamMask network exceeds the analytical mathematical approach by a 80% faster processing time.

Experimental verification of the measurement accuracy

The SiamMask network has been modified to process a camera image as direct input. For the experimental evaluation of the measurement accuracy of the network two CFTs were scanned with the camera. Here the CFT'S had a larger distance at the beginning than at the end. At 15 locations the distance between the CFT's was measured and automatically computed with the SiamMask network. Here $N = 48$ individual measurements were made for each location. In addition to the mean μ_{auto}, the standard deviation sd was also calculated to obtain the measurement uncertainty. For comparison, the gap between the tapes μ_{manual} was also measured manually with a caliper gauge. The maximum accuracy for the manual measurement is 0.05 mm. The relative error ε_r between manual and automatic measurement is calculated by

$$\varepsilon_r = \frac{\mu_{auto} - \mu_{manual}}{\mu_{manual}}. \tag{12}$$

Table 4 show that this resulted in a maximum error of 0.0806% with an average relative error of 0.0216%. The best measurement achieved a fit with a relative error of −0.0018%.

Assuming that all automatic measurements have approximately the same standard deviation, a pooled variance of 0.0422 mm is obtained. This indicates that the measurements with the network have higher accuracy than those with the caliper.

Table 4 The comparison between the manual and automatic measurement shows a maximum relative error of 0.0806%

μ_{manual} in mm	μ_{auto} in mm	$|\Delta|$ in mm	sd	ε_r (%)	p value	0-Hyp
0.25	0.256	0.006	0.0481	0.02583	0.46646	1
0.30	0.324	0.024	0.0204	0.08056	0.00000	0
0.50	0.511	0.011	0.0487	0.02167	0.19620	1
0.70	0.666	0.034	0.0385	−0.04792	0.00000	0
0.80	0.798	0.001	0.0519	−0.00182	0.90204	1
1.00	0.974	0.027	0.0248	−0.02687	0.00000	0
1.10	1.104	0.004	0.0180	0.00360	0.11179	1
1.30	1.277	0.022	0.0338	−0.01715	0.00017	0
1.40	1.359	0.041	0.0401	−0.02902	0.00000	0
1.60	1.564	0.035	0.0318	−0.02201	0.00000	0
1.70	1.689	0.011	0.0734	−0.00662	0.20342	1
1.90	1.926	0.026	0.0498	0.01349	0.00170	0
2.15	2.136	0.014	0.0512	−0.00640	0.08122	1
2.30	2.283	0.017	0.0353	−0.00752	0.00153	0
2.40	2.431	0.031	0.0322	0.01285	0.00000	0

The largest deviation of both measurements was 0.0410 mm. Thus the automatic measuring system shows sufficient accuracy for the production process

To investigate the statistical relationship between manual and automatic measurement a one-sided Wilcoxon test was applied to the difference between $\mu_{auto} - \mu_{manual}$, denoted by Δ.

The 0-hypothesis concerns whether the manual measurements match the automatic measurements from a statistical point of view. As can be identified in Table 1, the 0-hypothesis was rejected in several cases when $|\Delta|$ exceeds around 0.017 mm. The rejection of the 0-hypothesis is related to the lower measurement accuracy of the caliper gauge. Here the usually p-value limit of 5% was applied. In any case, however, $|\delta|$ is well above the production requirements where just an accuracy of approximately 0.2–0.3 mm can be achieved.

In summary, the automatic measurement exceeds the manual measurement in terms of accuracy and is therefore suitable for the intended AFP production process.

Conclusions and future work

It has been shown that a state of the art object segmentation networks like SiamMask can track and segment carbon fiber tapes during the layup process with sufficient accuracy. The network achieved around 80% faster calculation time, compared to the analytical-mathematical approach presented. Its performance is strongly hardware-dependent, but image processing and gap measurement under one second could be already achieved with a standard laptop without GPU support.

However, it is also shown that the accuracy of the gap measurement depends significantly on the segmentation masks which have currently only a size of 127×127 pixels. A modification of the network architecture, i.e. an increase of the mask resolution, offers great potential to increase the measurement precision. An equal image resolution of segmentation mask and input image would be ideal but is at the expense of computational effort and real-time capability.

In the current setup the total variance, with 0.0422 mm, of the automatic segmentation measurement is sufficient for the AFT process and the post-processing with 3D printing methods. This shows that segmentation methods, based on artificial neural networks have great potential to significantly improve complex manufacturing processes.

The present study also shows a method to generate training and testing data automatically. The generation of that data is based on recordings of the CFT placement during the machine operation. Here, it could be shown that the automatically generated data successfully trains the network. However, the practical application on the machine shows that complete virtual training can only be applied conditionally since errors in the segmentation generation cannot be excluded. A manual check of the training data is therefore recommended. Furthermore, segmentation errors can occur with changing light conditions. These errors can be minimized with larger training data sets .

In the existing network, measurements and inspections are performed based on the generated image segmentation. The comparison between the manual inspection, consisting of the gap measurement of two adjacent tapes, and the automatic evaluation shows that the SiamMask network generally achieves higher accuracy. The fast response time of the network enables online inspection, which allows error correction during tape placement.

Recent developments in Siamese networks also enable more reliable tracking decisions and the ability to retrace lost objects (Voigtlaender et al. 2020). It is conceivable to extend the network to provide quality features, machine control commands, and measurement values directly from the neural network without the need for explicit segmentation. The necessary training data is theoretically already available, since it is generated during the production process.

Future work can extend the processing of camera data and also analyze other process data that have an impact on material quality, such as melt temperature and feed rate (Brüning et al. 2017). This multi-causal approach enables a holistic view of the AFP process and is a step closer to smart manufacturing, as the machine can learn optimal process variables. The advantage of neural networks is that they can be arbitrarily extended with suitable training data while maintaining the single-system approach.

Declarations

References

Bahar, M., & Sinapius, M. (2020). Adaptive feeding roller with an integrated cutting system for automated fiber placement (AFP). *Journal of Composites Science*. https://doi.org/10.3390/jcs4030092.

Basu, M. (2002). Gaussian-based edge-detection methods—A survey. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews), 32*(3), 252–260.

Bertinetto, L., Valmadre, J., Henriques, J. F., Vedaldi, A., & Torr, P. H. (2016). Fully-convolutional Siamese networks for object tracking. In *European conference on computer vision* (pp. 850–865). Springer

Brüning, J., Denkena, B., Dittrich, M. A., & Hocke, T. (2017). Machine learning approach for optimization of automated fiber placement processes. *Procedia CIRP, 66,* 74–78.

Cemenska, J., Rudberg, T., & Henscheid, M. (2015). Automated in-process inspection system for AFP machines. *SAE International Journal of Aerospace, 8,* 303–309.

Ghayour, M., Hojjati, M., & Ganesan, R. (2020). Effect of tow gaps on impact strength of thin composite laminates made by automated fiber placement: Experimental and semi-analytical approaches. *Composite Structures, 248,* 112536.

Halbritter, A., & Harper, R. (2012). Big parts demand big changes to the fiber placement status quo. In *SME composites manufacturing*.

He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 770–778).

Heinecke, F. (2019). Strukturmechanische Auswirkung fertigungsbedingter Imperfektionen aus Faserverbundablegeprozessen

Heinecke, F., & Willberg, C. (2019). Manufacturing-induced imperfections in composite parts manufactured via automated fiber placement. *Journal of Composites Science, 3*(2), 56.

Ibarra-Castanedo, C., Piau, J. M., Guilbert, S., Avdelidis, N. P., Genest, M., Bendada, A., & Maldague, X. P. V. (2009). Comparative study of active thermography techniques for the nondestructive evaluation of honeycomb structures. In *Research in nondestructive evaluation*.

Kristan, M., Matas, J., Leonardis, A., Vojir, T., Pflugfelder, R., Fernandez, G., et al. (2016). A novel performance evaluation methodology for single-target trackers. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 38*(11), 2137–2155. https://doi.org/10.1109/TPAMI.2016.2516982.

Liu, D., Lohse-Busch, H., Toropov, V., Hühne, C., & Armani, U. (2016). Detailed design of a lattice composite fuselage structure by a mixed optimization method. *Engineering Optimization, 48*(10), 1707–1720.

Marr, D., & Hildreth, E. (1980). Theory of edge detection. *Proceedings of the Royal Society of London Series B Biological Sciences, 207*(1167), 187–217.

Meister, S., Wermes, M. A. M., Stüve, J., & Groves, R. M. (2020). Algorithm assessment for layup defect segmentation from laser line scan sensor based image data. In *Sensors and smart structures technologies for civil, mechanical, and aerospace systems 2020* (vol. 11379, p. 1137918). International Society for Optics and Photonics.

Minaee, S., Boykov, Y. Y., Porikli, F., Plaza, A. J., Kehtarnavaz, N., & Terzopoulos, D. (2021). Image segmentation using deep learning: A survey. *IEEE Transactions on Pattern Analysis And Machine Intelligence*.

Nguyen, D. C., Krombholz, C., & Röstermundt, D. (2012). Einfluss einer online bahnkorrektur auf die materialeigenschaften von prepreg tows im fiber placement prozess. Deutscher Luft- und Raumfahrtkongress.

Ondrašovič, M., & Tarábek, P. (2021). Siamese visual object tracking: A survey. *IEEE Access, 9,* 110149–110172.

Oromiehie, E., Chakladar, N., Rajan, G., & Prusty, B. (2019). Online monitoring and prediction of thermo-mechanics of AFP based thermoplastic composites. *Sensors, 19*(6), 1310. https://doi.org/10.3390/s19061310.

Parmar, H., Khan, T., Tucci, F., Umer, R., & Carlone, P. (2021). Advanced robotics and additive manufacturing of composites: Towards a new era in industry 4.0. In *Materials and manufacturing processes* (pp. 1–35).

Pieczynski, W., & Tebbache, A. N. (2000). Pairwise Markov random fields and segmentation of textured images. *Machine Graphics and Vision, 9*(3), 705–718.

Plath, N., Toussaint, M., & Nakajima, S. (2009). Multi-class image segmentation using conditional random fields and global classification. In *Proceedings of the 26th annual international conference on machine learning* (pp. 817–824).

Rakhshbahar, M., & Sinapius, M. (2018). A novel approach combination of automated fiber placement (AFP) and additive layer manufacturing (ALM). *Journal of Composite Sciences, 2,* 42.

Sacco, C., Radwan, A. B., Anderson, A., Harik, R., & Gregory, E. (2020). Machine learning in composites manufacturing: A case study of automated fiber placement inspection. *Composite Structures, 250,* 112514.

Sawicki, A. J., & Minguet, P. J. (1998). *The effect of intraply overlaps and gaps upon the compression strength of composite laminates*. American Institute of Aeronautics and Astronautics Inc.

Schmidt, C., Denkena, B., Hocke, T., & Völtzer, K. (2017). Influence of AFP process parameters on the temperature distribution used for thermal in-process monitoring. *Procedia CIRP, 66,* 68–73.

Schmitt, R., Pfeifer, T., & Orth, A. (2006). Feasible production of fibre-reinforced composites through inline inspection with machine vision. In *Proceedings of the IMEKO world congress, Rio de Janeiro, Brazil*.

Sezgin, M., & Sankur, B. (2004). Survey over image thresholding techniques and quantitative performance evaluation. *Journal of Electronic Imaging, 13*(1), 146–165.

Soucy, K. A. (1996). In-process monitoring for quality assurance of automated composite fabrication. In D. O. Thompson & D. E. Chimenti (Eds.), *Review of progress in quantitative nondestructive evaluation* (pp. 2225–2231). Berlin: Springer.

Sun, S., Han, Z., Fu, H., Jin, H., Dhupia, J. S., & Wang, Y. (2020). Defect characteristics and online detection techniques during manufacturing of FRPS using automated fiber placement: A review. *Polymers, 12*(6), 1337.

Tao, Y., Jia, S., Duan, Y., & Zhang, X. (2016). An online detection method for composite fibre tow placement accuracy. *Proceedings of the Institution of Mechanical Engineers, Part B: Journal of Engineering Manufacture, 230*(9), 1614–1621. https://doi.org/10.1177/0954405416640189.

Utkin, L., Kovalev, M., & Kasimov, E. (2021). An explanation method for Siamese neural networks. In *Proceedings of international scientific conference on telecommunications, computing and control* (pp. 219–230). Springer.

Voigtlaender, P., Luiten, J., Torr, P. H., & Leibe, B. (2020). Siam R-CNN: Visual tracking by re-detection. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition* (pp. 6578–6588).

Völtzer, K. (2018). Online-Prozessüberwachung von Automated Fiber Placement Prozessen auf Basis der Thermografie. Fakultät für Maschinenbau der Gottfried Wilhelm Leibniz Universität Hannover.

Wang, Q., Zhang, L., Bertinetto, L., Hu, W., & Torr, P. H. S. (2019). Fast online object tracking and segmentation: A unifying approach. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 1328–1338).

Wu, A. Y., Hong, T. H., & Rosenfeld, A. (1982). Threshold selection using quadtrees. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 1,* 90–94.

Zambal, S., Heindl, C., Eitzinger, C., & Scharinger, J. (2019). End-to-end defect detection in automated fiber placement based on artificially generated data. In *Fourteenth international conference on quality control by artificial vision* (vol. 11172, p. 111721G). International Society for Optics and Photonics.

Zhang, D., Zheng, Z., Wang, T., & He, Y. (2020). HROM: Learning high-resolution representation and object-aware masks for visual object tracking. *Sensors, 20*(17), 4807.

Zhu, J., Zhang, G., Zhou, S., & Li, K. (2021). Relation-aware Siamese region proposal network for visual object tracking. *Multimedia Tools and Applications, 80*(10), 15469–15485.