



# Deep convolutional neural networks: Outperforming established algorithms in the evaluation of industrial optical coherence tomography (OCT) images of pharmaceutical coatings

Matthias Wolfgang<sup>a</sup>, Michael Weißensteiner<sup>a</sup>, Phillip Clarke<sup>a</sup>, Wen-Kai Hsiao<sup>a</sup>, Johannes G. Khinast<sup>a,b,\*</sup>

<sup>a</sup> Research Center Pharmaceutical Engineering GmbH, Graz, Austria

<sup>b</sup> Institute for Process and Particle Engineering, Graz University of Technology, Graz, Austria

## ARTICLE INFO

### Keywords:

Convolutional neural networks (CNNs)  
Optical coherence tomography (OCT)  
Image segmentation  
Coating layer thickness  
Single- and multi-layered coatings

## ABSTRACT

This paper presents a novel evaluation approach for optical coherence tomography (OCT) image analysis of pharmaceutical solid dosage forms based on deep convolutional neural networks (CNNs). As a proof of concept, CNNs were applied to image data from both, in- and at-line OCT implementations, monitoring film-coated tablets as well as single- and multi-layered pellets. CNN results were compared against results from established algorithms based on ellipse-fitting, as well as to human-annotated ground truth data. Performance benchmarks used include, efficiency (computation speed), sensitivity (number of detections from a defined test set) and accuracy (deviation from the reference method). The results were validated by comparing the output of several algorithms to data manually annotated by human experts and microscopy images of cross-sectional cuts of the same dosage forms as a reference method. In order to guarantee comparability for all results, the algorithms were executed on the same hardware. Since modern OCT systems must operate under real-time conditions in order to be implemented in-line into manufacturing lines, the necessary steps are discussed on how to achieve this goal without sacrificing the algorithmic performance and how to tailor a deep CNN to cope with the high amount of image noise and alterations in object appearance. The developed deep learning approach outperforms static algorithms currently available in pharma applications with respect to performance benchmarks, and represents the next level in real time evaluation of challenging industrial OCT image data.

## 1. Introduction

Optical coherence tomography (OCT) is an established imaging technology in the fields of medicine (de Boer et al., 2017), art conservation (Rouba et al., 2008) and non-destructive testing (Nemeth et al., 2013). In our past work we have established the first commercial applications of OCT in the pharmaceutical industry as a contactless and real-time process analytical technology (PAT) for solid dosage form coating processes. Specific aspects of industrial pharmaceutical film coatings that can be investigated via OCT have been discussed in the literature, e.g., film thickness and variability (Markl et al., 2015a), surface roughness (Markl et al., 2018) and film homogeneity (Sacher et al., 2019).

Zeitler and Gladden were the first ones to propose OCT for non-destructive investigations of pharmaceutical coatings (Zeitler and

Gladden, 2009). At that time, OCT images could only be evaluated manually via image manipulation software or Matlab. In 2010, Zhong et al. compared the merits of OCT and Terahertz pulsed imaging (TPI) for the examination of pharmaceutical coating layers on tablets. They applied the phase shifting method for OCT image reconstruction to achieve very high axial resolution for the first time on pharmaceutical coatings (Zhong et al., 2010). Lin et al. presented an automated interface detection algorithm for pharmaceutical tablet coatings in 2015 based on Wavelet denoising (Lin et al., 2015), which was then further refined and applied on in-line acquired data (Lin et al., 2016).

Also in 2015, another automatable method for image evaluation of pharmaceutical coatings based on a circular fit algorithm (Markl et al., 2015a) was presented, under which OCT images were evaluated based on single interferograms (analogous to a sonography “A-scan”) by extracting gradient features via one-dimensional convolutional kernels.

\* Corresponding author at: Institute for Process and Particle Engineering, Graz University of Technology, Graz, Austria.

E-mail address: [khinast@tugraz.at](mailto:khinast@tugraz.at) (J.G. Khinast).

<https://doi.org/10.1016/j.ijpx.2020.100058>

Based on the extracted features, the coating interfaces were determined using a shortest path algorithm (Markl et al., 2015b). This approach worked well for tablets of a defined curvature, provided that the OCT images are of adequate contrast and resolution. Furthermore, the expected curvature had to be pre-set manually for this type of algorithm. Object features deviating from the predefined curvature (e.g. the dosage form, edges and embossing) could lead to inaccurate results or even complete detection failure. The 2018 review paper from Lin et al. provides a concise overview of what has been achieved in terms of the application of OCT on pharmaceutical coatings (Lin et al., 2018) up until this time.

Following the automated circle-fit algorithm, the first real-time algorithm method based on elliptical fits was developed, which is capable of extracting features with fewer pre-set parameters and at considerably higher speeds in industrial in-line configurations (Sacher et al., 2019). This method allows to analyze 2-dimensional gradient features of full OCT images (analogous to sonography "B-scans") calculated via elliptical kernels (2-dimensional anisotropic kernels), resulting in significantly more robust features for the elliptical fit optimization. The derived feature map serves as a basis for the extraction and evaluation of the actual image information. In contrast to the circle-fit approach, the relative position of a tablet tomogram inside the B-scan plays almost no role in the automated layer detection. However, this elliptical fit method has limitations. First, the edges, holes and highly porous layers cannot be reliably detected. Second, extreme inclinations of layers in the image and interfaces mirrored on the top of the OCT image cannot be evaluated. Furthermore, it is important to note, that only OCT images of sufficient image contrast are evaluable, regardless of the algorithms used. This implies that the limitations of this imaging technology are associated with the scattering properties of the examined coating, for example, coatings containing high amounts of scattering pigments like titanium dioxide cannot be evaluated, as summarized in (Lin et al., 2017).

Our previous work focused mainly on comparing OCT evaluation to other indirect coating measurement methods (Wahl et al., 2019) and validating the (ellipse fit based) automated real-time algorithms for tablet film coating processes (Wolfgang et al., 2019). However, an image analysis method with a high degree of flexibility, efficiency and robustness is still required since industrial OCT images are unpredictable in terms of product shape, relative position and visual appearance. With that regard, current machine learning applications in the field of computer vision may provide superior solutions. Unlike in other fields of OCT (e.g., retinal pathology investigation), tomographic images of pharmaceutical coatings are very challenging for machine learning due to a high and variable level of image noise, varying contrast in the course of long runs, an often dramatic change in the layer appearance over the process (layer growth in the course of a coating run) and dynamic distortions caused by arbitrary movement of the target products.

Of the currently available machine learning approaches, convolutional neural networks (CNN) appear to be the most promising in this application, especially fully convolutional networks (FCN) which can analyze input image data of varying sizes (Long et al., 2015). These models have shown state-of-the-art results in semantic image segmentation in various fields, e.g., computer vision systems for autonomously driving cars (Zhang et al., 2018) and medical image analysis (Ronneberger et al., 2015). Methods of identifying retinal tissue and intra-fluid layers via a fully convolutional architecture called "ReLayNet", as proposed by Roy et al., have been successfully applied in the field of retinal OCT image segmentation (Roy et al., 2017) using a U-Net architecture.

Recent studies, applying CNNs to evaluate OCT image data of scattering layers using a U-Net architecture, investigated the merits of multi-layer segmentation by introducing "dense connected blocks" with residual connections. These blocks increased the accuracy of outputs, yet also the required calculation effort (Kepp et al., 2019). Another approach combined CNN segmentation based on U-Net, with extensive post-processing by Savitzky-Golay-filtering for interface detection and

Fourier domain filtering for smoothing of contours, in order to segment follicular structures of skin (del Amor et al., 2020).

However, since such deep neural networks require a large amount of computational power for real-time results, one of the challenges is to develop a lean network architecture while maintaining an accurate and robust performance. Another challenge is to provide qualitative training and validation data. Since industrial OCT images of pharmaceutical coatings have very high variations in appearance, contrast and sometimes shapes, a tailor-made neural network, based on the U-Net architecture introduced by Roy et al., was developed and trained. An appropriate training data set for the OCT application should include a large number of high-quality images, with a large variation in terms of visual appearance, manually segmented by a human expert. As in other machine learning applications, the accuracy of annotated training data is critical for the performance of the resulting model.

Given that established static algorithms with fixed parameters perform well when the OCT images of tablets are correctly presented, we compared various challenging scenarios with the established static algorithms to the CNN models. Especially tricky images include those with poor contrast, high image noise, arbitrary shapes and a highly scattering or porous layer appearance.

In this study, we present a novel approach, which applies fully convolutional deep learning algorithms to live OCT imaging of non-static tablets or pellets for process monitoring. Machine learning alternatives, such as classical convolutional networks, random forests and other pixel classifications approaches were discarded due to their extensive calculation demands, which obviate the real-time capability requirement.

## 2. Materials

In order to generate OCT image data of challenging nature, several established solid dosages forms for acquiring suitable OCT images were selected. The selected products are summarized in Table 1.

All tested products are commercial samples, with the exception of Thrombo ASS tablets. These were coated in-house using various process parameters and coating compositions in order to generate different appearances of the final layers due to the varying process dynamics. The goal was to investigate the robustness of the developed algorithms in the presence of process variations.

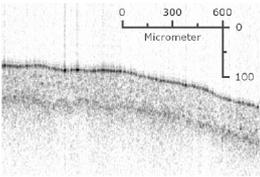
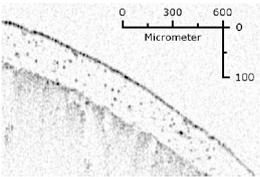
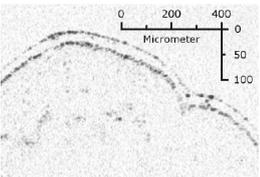
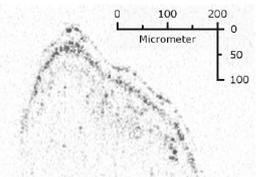
### 2.1. Tablets tested

#### 2.1.1. Scattering tablet coatings

Thrombo ASS 100 mg tablets were chosen as an example of a strongly scattering enteric coating, containing more than 30% of scattering talcum particles. In order to generate OCT in-line data of the coating trials for this product, original tablet cores (GL Pharma, Lanach, Austria) were coated in a DRIACONTI-T pharma lab coater (Driam Anlagenbau GmbH, Eriskirch, Germany). Coating was performed in a semi-continuous manner in three subsequent chambers, with a filling volume of 1.8 l and 1.4 kg of tablets in each run. Each chamber has a Schlick spray nozzle (Düsen-Schlick GmbH, Coburg, Germany) with a bore diameter of 0.5 mm. Atomizing pressure of 0.7 bar and a drum speed of 22 rpm remained constant throughout all experiments. The individual process settings for the coating runs are listed in Table 2.

The coating formulation consists of EUDRAGIT L 30 D-55, a common enteric coating prepared according to the manufacturer's specification (Evonik Industries AG, 2011). Iron (III) oxide with the percentage indicated by the run names (0.5%, 2% and 10% weight) was added to the basic composition to generate highly scattering coating layers. Two additional runs were done with 0.5% and 2.5% weight of indigo carmine as a blue dye to test the dye's impact on the OCT images. One run was performed with half the amount of the talcum in the coating solution. In all remaining runs (indicated only by number in Table 2), coating solutions as specified by the manufacturer was applied. These runs were

**Table 1**  
Overview of products for generating OCT images.

	Scattering	Porous	Arbitrary Shaped	Multi-layered
Product name	Thrombo ASS 100 mg	Glucotrol XL 5 and 10 mg	Effexor XR	Detrol
Type	Biconvex tablets	Biconvex tablets	Pellets	Pellets
Coating	Enteric coating	Osmotic coating	Extended release coating	Extended release coating
OCT image				

**Table 2**  
Process parameters of the examined in-line runs to achieve varying appearances for the resulting OCT coating layer images.

Run name	Solution applied [g]	Duration [min]	Flow rate [g/min/kg]	T <sub>Inlet</sub> [°C]	P <sub>Pattern</sub> [bar]
0.5% Fe <sub>2</sub> O <sub>3</sub>	681	112	4.5	52.5	1.3
2% Fe <sub>2</sub> O <sub>3</sub>	641	112	4.5	52.5	1.3
10% Fe <sub>2</sub> O <sub>3</sub>	680	112	4.5	52.5	1.3
0.5% Dye	677	112	4.5	52.5	1.3
2% Dye	643	112	4.5	52.5	1.3
2.5% Dye	618	112	4.5	52.5	1.3
15% Dye	789	201	3	45	1
Talcum					
Run 1	680	167	3	45	1
Run 2	780	84	6	45	1
Run 3	675	167	3	60	1
Run 4	710	84	6	60	1
Run 5	618	167	3	45	1.6
Run 6	642	84	6	45	1.6
Run 7	727	112	4.5	52.5	1.3
Run 8	637	112	4.5	52.5	1.3
Run 9	681	112	4.5	52.5	1.3
Low Temp.	640	113	3	35	1

performed using different process parameters to vary the OCT image layer appearance based only on the process conditions. A detailed description of all experiments conducted has been summarized in [Zettl, 2017](#).

OCT in-line monitoring was conducted using an OCT system developed in-house as described in [Markl et al., 2014](#), with a 1D OCT sensor mounted inside the third coater chamber scanning through the perforation of the pan. The OCT system features an axial resolution of 4.1  $\mu\text{m}$  (in air) based on the light source applied (832 nm central wavelength and 75 nm bandwidth) and a lateral resolution of 14  $\mu\text{m}$  (in focus) determined by the optics in the sensor head. The maximum A-scan rate of the system is 59.2 kHz, resulting in a maximum acquisition speed of 57.8 frames per second with  $1024 \times 1024$  pixel (12 bit) OCT image resolution (B-scan). To ensure comparability, data evaluation for all experiments in this study was conducted on the same hardware after the various trials.

Validation of the algorithms for Thrombo ASS OCT images was undertaken by comparing the results of a predefined test set to manually annotated data generated by human experts and to a reference method. The latter was data generated by evaluating the light microscopy images of cross-sectional cuts of the tablets examined via OCT, as described in [Wolfgang et al., 2019](#).

### 2.1.2. Porous tablet coatings

Commercial Glucotrol XL 5 mg and Glucotrol XL 10 mg tablets were used as examples of a porous coating with a porous asymmetric semi-permeable membrane system to generate an osmotic push-pull system in

combination with the tablet core ([Donald L. Wise, 2000](#)). The tablet samples were examined using a commercial at-line sampling device, with the corresponding OCT system "OSeeT Pharma 1D" (both from Phyllon GmbH, Graz, Austria). The tablet samples consisting of approx. 100 tablets each were moved into a small perforated bench-top drum setup, mimicking the behavior of a tablet bed inside a coating drum. The OSeeT Pharma 1D system has the same axial resolution of 4.1  $\mu\text{m}$  (in air) at a lateral resolution of 14  $\mu\text{m}$  but at a higher acquisition speed of 100,000 A-scans per second.

Validation of the results of Glucotrol XL for the various algorithms was conducted by comparing them to the evaluation of the same OCT data set performed by two human experts using the scientific open source image processing software ImageJ2 ([Rueden et al., 2017](#)).

### 2.2. Pellets tested

As an example of arbitrarily-shaped single layer pellets, the commercial product Effexor XR (Venlafaxin) was chosen, with the pellet cores containing the drug and coated with an extended release coating layer. In order to assess multi-layered pellet coatings, the commercial product Detrol (Tolterodine tartrate) was selected, which contains multi-layered coated pellet cores, with layers containing the drug and layers moderating the release behavior. All pellet samples were provided by Pfizer Inc.

All pellets were examined in an at-line setup, using a rotating Petri dish Schütt Petriturn-E lab tool (Schuett-biotec, Göttingen, Germany), combined with a commercial in-line 1D OCT sensor (Phyllon GmbH, Graz, Austria) mounted above it, scanning the pellets as they passed by the sensor. The sensor was connected to the corresponding Phyllon OSeeT Pharma 1D base-unit for data acquisition.

Validation of the layer segmentation for the CNN algorithms was visually performed by four human experts based on the appearance of the OCT input and output images.

### 2.3. Computational hardware

To ensure comparability of the results, all evaluations in this study were executed on the same hardware. All training, evaluation and validation steps were performed using a NVidia GTX 1080 Ti graphics card (GPU), which is an integral part of the commercial OSeeT Pharma 1D system. The CPU used was an Intel Core i5 with 8 GB of DDR3 RAM. The operating system was Microsoft Windows 10.

## 3. Methodology

In order to benchmark the performance, the CNNs developed were compared to the most evolved, in-line version of the ellipse-fit approach as described in [Sacher et al., 2019](#).

### 3.1. Training data generation and inference

One of the most challenging tasks in generating CNNs is the preparation of a suitable amount of training data (several hundred annotated images). This involves manually annotating entire OCT images at the pixel level. For this, experienced experts attempted to identify the coating layer area as accurately as possible to produce a binary segmentation mask, which serves as ground-truth information describing foreground (coating) and background pixels. The annotated image data was split into training data (approximately 80% of all images annotated) and validation data, in order to generate a dice score. External validation of the results was completed based on OCT data, which had been validated by microscopy results of cross-sectional cuts at the same positions of OCT investigation. Unlike more established fields of application where CNNs are used, no publicly available databases for annotated OCT images of pharmaceutical dosage forms currently exist. As a consequence, all annotation data, as well as validation data, had to be generated from scratch for this study.

During CNN training, the network loss is calculated based on these segmentation masks and back propagated throughout the network to optimize the layer weights accordingly. The network loss describes the network error in the training data during training and is optimized as close to zero as possible for training and validation data in an iterative manner. In this way, the model learns how foreground and background pixels appear with respect to their neighborhood.

Fig. 1 provides an example of an OCT image that contains three tablet instances of the product Glucotrol XL 5 mg and its ground-truth segmentation mask as binary images. The right image shows the corresponding inference output of a trained CNN model as an overlay for the illustration. This output segmentation map represents the probability of each pixel being a foreground or “coating-layer” pixel, which results in an almost full segmentation of all visible coating layers, although of varying intensity and orientation.

As the image patch detail in Fig. 1 indicates, a full coating layer segmentation of unseen OCT image data is possible under this approach, regardless of the elliptical product shape.

To further extend the amount of available OCT image training data in terms of quantity and variety, and to reduce the effort of manual image annotation, the datasets were augmented by random horizontal flipping

and additional random brightness variation during training. The reason why the images were not flipped vertically or rotated during augmentation is that the aim was to preserve the nature of an A-scan, including such effects as the OCT-signal sensitivity roll-off (de Boer et al., 2017). This secures the behavior and appearance of any optical reflection on the interfaces that the network should implicitly learn.

To overcome the potential lack of training image patch diversity in terms of unseen layer thicknesses, synthetic image training patches of  $128 \times 128$  pixel were additionally generated using the concept of Deep Convolutional Generative Adversarial Networks (DCGANs) (Radford et al., 2016). Neff et al. investigated an application of GANs to thorax X-ray images for medical image segmentation using a U-Net architecture (Neff et al., 2017). They showed that GANs have significant potential in terms of medical image training data synthesis, especially when datasets are small. This inspired the application of this data augmentation approach to coating thickness variation. During a real coating process, a CNN’s robustness against thickness is crucial. To investigate this approach, real images of thin and thick coating layers and synthetic images of in-between coating thicknesses were fed in the CNN during training by interpolating the GAN input vector. The goal was to mimic unseen coating thicknesses in order to improve the overall segmentation performance throughout the coating process. However, these experiments did not improve the layer segmentation accuracy, which is in agreement with the findings of Neff et al.

### 3.2. Fully convolutional neural network

The chosen CNN is based on the U-Net architecture that was first proposed by Ronneberger in 2015 (Ronneberger et al., 2015). This type of architecture is fully convolutional (Long et al., 2015) and consists of an encoder path and a decoder path, with additional skip connections in between. In total, the network has 23 trainable layers, hence it is a deep learning network. The encoder path has several subsequent 2D convolutional layers, each followed by a max-pooling layer. This decreases the spatial feature size, increasing the level of abstraction. In this way, relevant higher-level features can be encoded to represent contextual information. The subsequent decoder path up-samples the feature channels via 2D up-convolutions, whereby the context information is propagated up to the output layer (output segmentation map) by

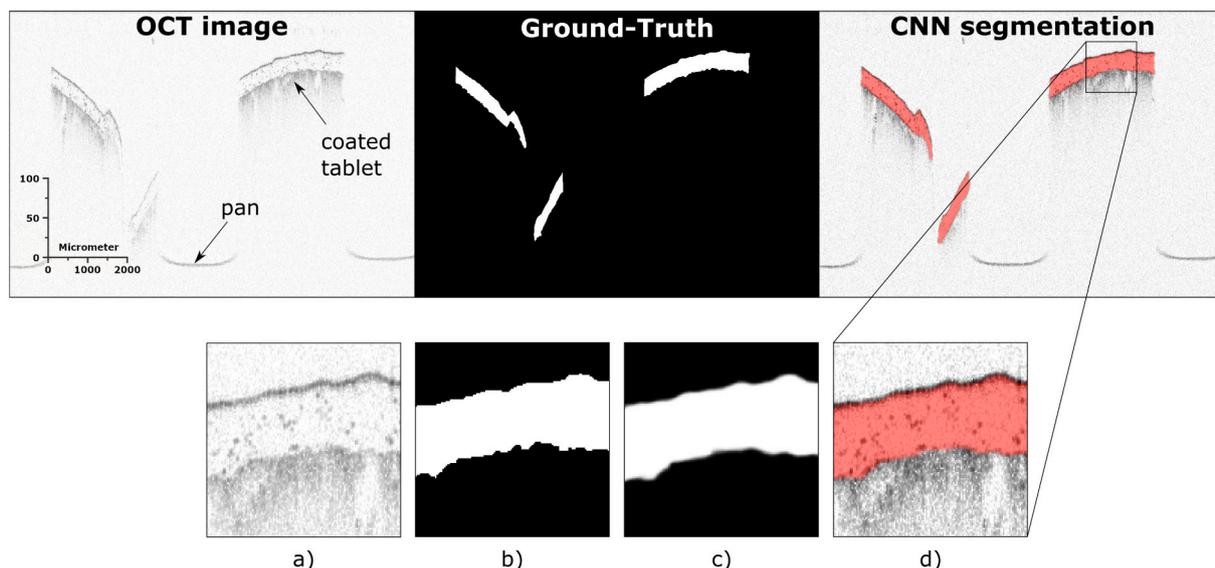


Fig. 1. Example of an OCT image for Glucotrol XL 5 mg product samples. Top-left: OCT image containing three scanned products arbitrarily positioned and orientated. Top-middle: Manually produced ground-truth segmentation mask with annotated coating layer pixels as a binary image. Top-right: CNN-based coating layer segmentation map overlay in red colour. Details of a cropped image patch (black bounding box): a) Cropped OCT image patch of  $128 \times 128$  pixel; b) ground-truth mask; c) CNN output segmentation map; and d) segmentation map overlay in red. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

additionally using the relevant feature channels from the encoder path via the skip connections. As shown by Ronneberger, this improves the overall segmentation performance since the finely-resolved low-level features can be used to refine the up-sampling of deeper, coarsely-resolved semantic information on the decoder path (Ronneberger et al., 2015). The working principle and the building blocks of this architecture are illustrated in Fig. 2.

Due to the inherent nature of 2D convolution used by the original U-Net architecture, the spatial dimensions are reduced after each convolutional layer. This would inevitably require cropping of the feature maps from the encoder path to fit the decoder feature dimensions via the skip connections. Using the original U-Net architecture would result in a lower resolution of the output segmentation map. In contrast, the proposed architecture retains the spatial size on each level of abstraction (each convolutional layer indicated by blue arrows in Fig. 2) by using zero-padding before each convolution as implemented in the Keras framework (Chollet et al., 2015). Only the max-pooling layers (indicated by turquoise arrows in Fig. 2) reduce the dimensions accordingly. As a result, all cropping operations were removed to obtain an output segmentation map of the same size as the input image. The second major difference from the U-Net architecture is that the developed CNN requires fewer output channels across all convolutional layers, providing an architecture that has approximately 60 times fewer parameters to train (reduction from 31,030,593 to 485,673 parameters in total), which correspondingly reduces the inference time and being approximately 28 times more efficient during inference. The speed-up factor is related to applying the original U-Net architecture, which provides an inference time of 455 ms per full OCT image on the same hardware. Our reduced model complexity, in contrast, shows an inference time of only 16 ms. This reduces the GPU memory consumption and satisfies the real-time requirement for the system.

The major advantage of this fully convolutional architecture is the lack of fully connected or dense layers, which provides a very convenient feature: If the input image size is changed, only the spatial dimensions of the intermediate results (the feature channels) are altered without having to adapt the network architecture to the actual input size. Considering that training images are cumbersome to generate, this

significantly boosts the number of available training samples. Training can be conducted on image patches of a smaller size rather than full-size images. The only requirement for these cropped training patches is to display a sufficiently large area that captures all relevant image features. For example, the network must “see” almost the full height of a thicker coating layer in a cropped patch, as shown in Fig. 1. It will become more reliable during training since it learns that two interfaces denote the coating layer. The textural information for the coating layer itself is also learned, which is an essential contribution to the network confidence upon application.

In summary, the above steps produce a fully convolutional architecture highly suitable for the intended task. Many more training samples can be provided (thousands of image patches) to the network, with only a limited number of full-size training images (a few hundred) available. However, arbitrarily large images may be fed in the trained image segmentation model during testing, which is only limited by the available GPU memory and possible real-time constraints.

### 3.3. CNN training

The basic training approach for all CNN models in this study uses randomly cropped training patches of reduced image size as input data. The 1:1 sampling ratio remains between the foreground and background patches to ensure balanced training, whereby a foreground image patch is defined as the one containing at least one foreground ground-truth pixel. The mini-batch size is set to 512 random image patches sampled from four random OCT images of the training set. As a measure of the network’s loss, the dice score is used, which is a score for overlap between two areas defined as:

$$Dice(P, M) = \frac{2|P \cap M|}{|P| + |M|}$$

where  $P$  is the predicted segmentation map and  $M$  is the ground-truth mask, both with pixels within the range [0,1]. Thus, the network applies a sigmoid output layer. A perfect overlap has a dice score of 1.0, whereas 0.0 describes disjointed areas.

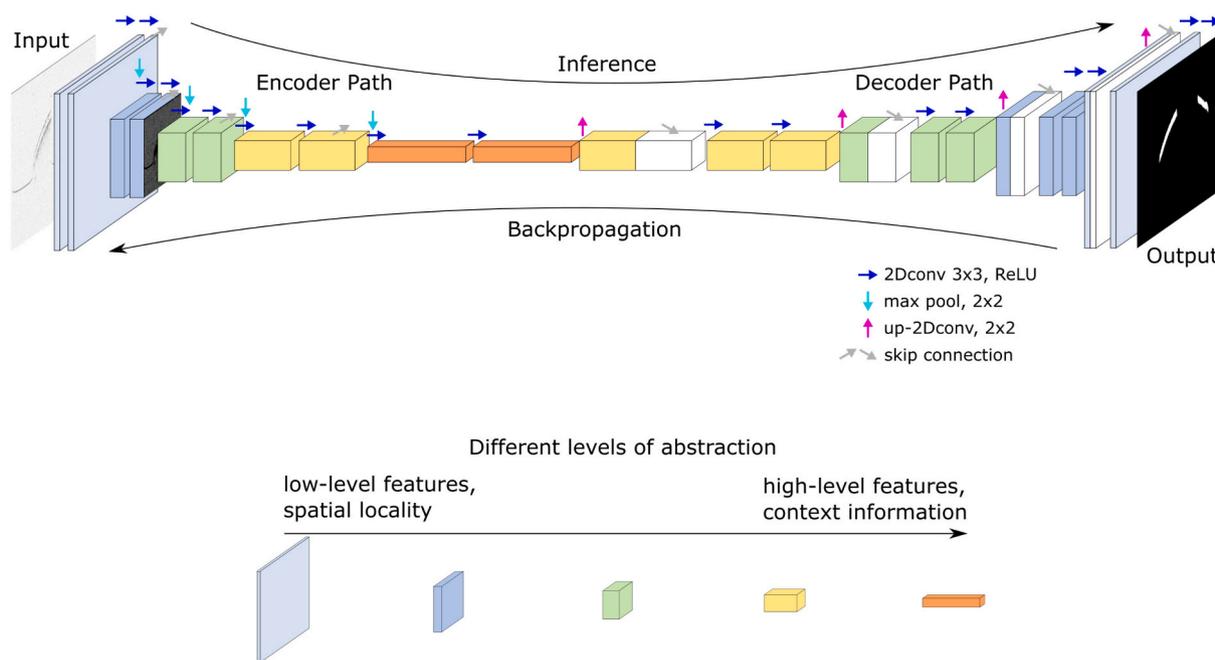


Fig. 2. The proposed CNN architecture schema described by an encoder and decoder path with skip connections in between (grey arrows). The input layer represents a grayscale OCT image. The output layer produces a segmentation map of pharmaceutical coating. This approach preserves the spatial dimensions after every convolution due to zero-padding. Intermediate data representations are illustrated for the encoder path, which increases the level of abstraction and the number of feature channels.

The network weight initialization was determined empirically and yielded the best results when the “he\_normal” initializer (He et al., 2015) was applied. The latter samples the weights according to a normal distribution with a modified standard deviation that considers the number of input neurons for each input layer. To learn the network parameters, the Adam optimizer (Kingma and Ba, 2015) was used as provided by the Keras framework (Chollet et al., 2015), which has empirically shown to be the most suitable and effective to train our network. To avoid overfitting, early-stopping was applied by focusing on the validation loss, terminating optimization of the training process when the underlying (validation) loss increased.

The same general approach may be applied to generate networks that can detect multiple features at once, e.g., multiple layers as part of the coating. Nevertheless, it requires additional output channels or output segmentation maps describing a specific layer. Segmentation is performed concurrently in the output layer, making the results for all coating layers available simultaneously.

In the course of this study, several CNNs were trained, each to serve a specific purpose. A specific process shown in Fig. 3 was followed. Once a small number of training images was generated (OCT image - segmentation mask pair), the network was trained and subsequently tested using a blind test set of approximately 10% of unseen training images (subset of the training dataset) to assess the validation loss. If the network performed well in terms of visual output and a dice score above 0.9 was achieved, training was stopped. Otherwise, more training data was added, some architecture details were improved or training parameters were enhanced. The final generic CNN model required 70 Thrombo ASS images and 45 Glucotrol XL 5 mg images for training. All implementations were written in Python 3.6 (Python Software Foundation, <https://www.python.org/>).

Our next task was to establish how many OCT images are required for learning a specific tablet type from scratch and what impact this has on the network generalization. To that end, the network was trained on an increasing dataset size of randomly sampled Thrombo ASS OCT images chosen from a fully annotated dataset of 83 full sized OCT images. The chart in Fig. 4 shows the segmentation performance for 10 fixed,

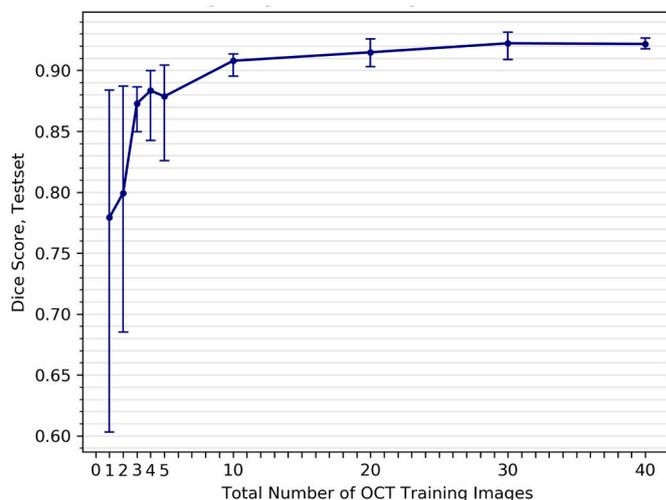


Fig. 4. CNN segmentation performance based on varying training set sizes for the Thrombo ASS product. Points show the average dice score over 5 runs for 10 unseen test images; Whiskers denote the maximum and minimum dice scores.

unseen images over an increasing training dataset size. The statistics are based on five training runs. Note that each additional training image potentially provides thousands of training image patches for training the network.

#### 4. Results

##### 4.1. Validation of the tablet and pellet results based on validations sets generated by experts

Validation of the results was performed based on test sets containing OCT images of coating layers of different appearance: single and multi-layered coatings. The OCT images were evaluated by human experts,

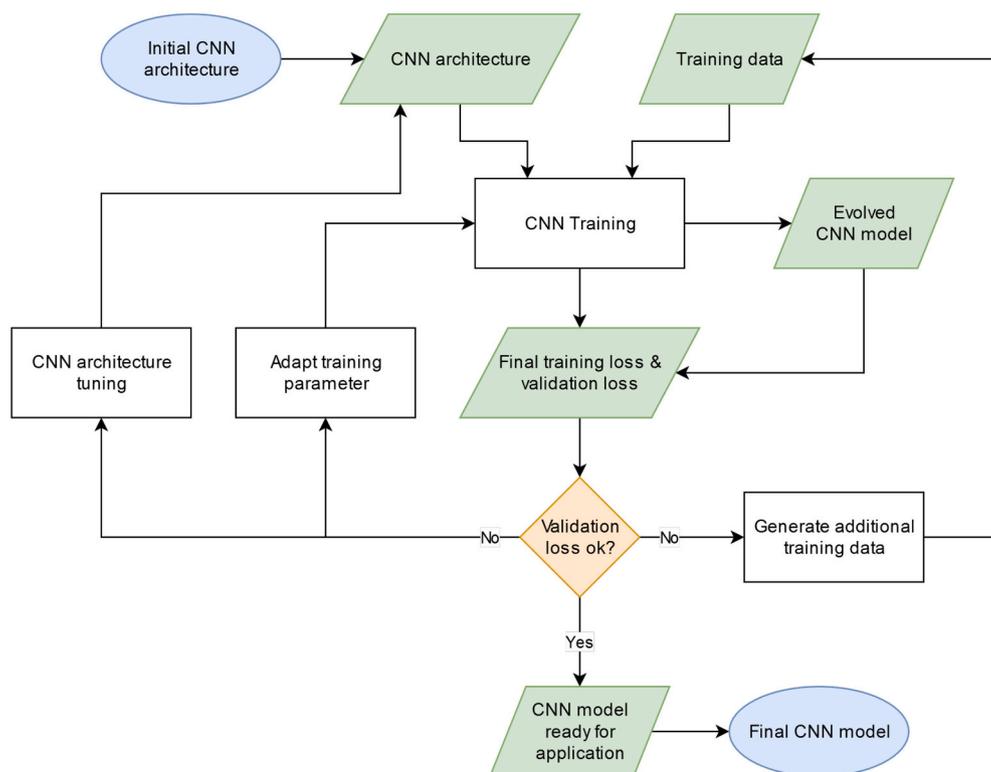


Fig. 3. Flowchart overview of how to iteratively evolve a CNN model. After an initial development of a CNN architecture, the CNN is trained on a given training dataset. If the final validation loss is sufficiently low, the model is considered to be ready for application. Otherwise the model is retrained and evaluated after either (i) tuning the CNN architecture, (ii) adapting the training parameters or (iii) generating additional training data, which ideally improves the CNN model performance.

forming the reference for validation. The findings were compared to the output of the corresponding neural networks. A summary is provided in Table 3, clearly indicating a strong agreement between ground truth and results for the CNN.

The dice score presented in Table 3 relates to the test images and not to training of the network. With respect to these results, the score measures the degree of variability of the images compared to the coverage of the network, which is in fact higher for arbitrarily shaped objects (e.g., pellets). The root mean squared error (RMSE) for the interfaces describes the deviation of each calculated interface from the “real” (reference) interface path.

The second validation approach compares the results for the automatically evaluated OCT images of Thrombo ASS tablets to the interface thicknesses and standard deviation in the coating thickness measured by means of microscopy as a reference. The comparison is summarized in Table 4. A detailed description of the procedure can be found in Wolfgang et al., 2019. For better comparison, the CNN output (a pixel area) was converted into a coating thickness value by dividing the segmented area through the number of contributing A-Scans and multiplying it by an OCT-system-specific conversion factor and a material-specific correction of the refractive index. For the Thrombo ASS, an assumed refractive index of the coating of 1.48 is used, as reported in Wolfgang et al., 2019. Mean values and standard deviations are calculated over all contributing A-Scans of the CNN output in order to be comparable.

#### 4.2. Comparison of computational speed

For the coated Thrombo ASS tablets, a broad range of OCT image appearances is recognized, which is reflected in a different performance under both CNN and ellipse-fit evaluation approaches, as shown in Fig. 6. A qualitative summary on the various OCT image features and their impact on the investigated algorithms is listed in Table 5. Especially challenging are those images, where the coating-core interface is faded or blurred and hardly visible, as it is the case when scattering pigments (e.g.,  $\text{Fe}_2\text{O}_3$ ) are present in the coating or the coating is highly porous or agglomerated, resulting in extensive scattering as well. However, it has to be pointed out that even for human experts it is nearly impossible to distinguish an interface for those images.

In order to compare the computational speed of the algorithms, the total time required to evaluate all OCT images of entire coating experiments was determined for all available test sets. Evaluation results for all 17 conducted coating experiments are summarized in Fig. 5. Note that raw data in these evaluations contain datasets of various sizes (3886–8734 full OCT images). Not all images contain evaluable information, i.e. yielding valid detections. The total time reported includes the time required for data reading from the hard drive and image analysis. Therefore, in a real-world in-line monitoring application, the

**Table 3**

Validation results for CNN and ellipse-fit compared to those obtained by human experts for test sets of OCT images. The results are on a pixel basis.

Sample name	Thrombo ASS 100 mg	Glucotrol XL 5 mg	Detrol	Effexor XR
Dosage Form	Tablet (enteric coat)	Tablet (osmotic coat)	Pellet (multi- layer)	Pellet (single- layer)
# Full Train Images	59	36	39	11
Coating Height (Reference), [px]	37.2 ± 3.1	45.8 ± 4.2	7.9 ± 1.7	15.3 ± 2.6
Coating Height (Prediction), [px]	36.4 ± 2.6	45.4 ± 3.0	8.0 ± 1.6	14.8 ± 2.1
Dice Score	0.93	0.94	0.67	0.86
RMSE Top Interface, [px]	1.5	1.7	1.5	1.8
RMSE Bottom Interface, [px]	3.1	3.5	1.6	2.5

**Table 4**

Comparison of results for CNN evaluation and light microscopy. Reported values are mean values ± standard deviation.

Sample	Light microscopy (µm)	CNN (µm)	Ellipse (µm)
1	75.2 ± 2.2	75.8 ± 3.5	76.5 ± 3.4
2	57.3 ± 3.8	59.2 ± 5.5	56.3 ± 3.3
3	74.4 ± 3.1	71.5 ± 4.7	73.3 ± 3.1
4	75.3 ± 6.1	80.0 ± 4.8	86.9 ± 3.6
5	65.1 ± 3.8	64.3 ± 5.1	67.3 ± 5.1
6	68.1 ± 3.6	67.5 ± 3.4	68.8 ± 4.1
7	57.8 ± 3.9	60.0 ± 3.6	62.7 ± 3.8
8	75.5 ± 8.7	66.2 ± 4.0	69.1 ± 3.9
9	64.4 ± 4.9	67.4 ± 2.5	68.6 ± 4.5
10	80.1 ± 4.9	79.1 ± 3.9	84.0 ± 3.3
11	54.1 ± 5.3	59.0 ± 3.5	59.5 ± 3.7
Mean <sub>all</sub>	68.0 ± 8.4	68.2 ± 7.3	70.3 ± 3.8

**Table 5**

Qualitative assessment of image features and their impact on the performance of the algorithms.

Image feature	Range of variation	CNN able to evaluate the whole range of variation	Ellipse-fit able to evaluate the whole range of variation
Scattering properties of coating layer	From weak to strong scattering	Yes, but challenging for strongly scattering layers	No, strongly scattering layers are impossible to evaluate
Definition of coating-core interface	Well defined to nearly invisible for humans	Yes, but challenging for faded/blurred interfaces	No, weak interfaces are not properly detected
Fringes and other image artefacts	None to a few	Yes, no impact on performance	Yes, but impact on proper interface detection
Surface roughness	Smooth to well visible topography	Yes, no impact on performance	Yes, but impact on proper interface detection
Orientation of objects in image	Perpendicular to high inclined	Yes, object orientation does not affect the segmentation	No, highly inclined objects are not properly evaluated

evaluation time required could be significantly lower since in this case raw data is transported directly from the sensor to the GPU via the PCI express interface.

The tested CNN is slightly more efficient in terms of computational speed than the ellipse-fit evaluation. However, as stated above, the major part of elapsed time refers to read times from the hard drive for both approaches. Furthermore, the CNN-based evaluation is capable of detecting multiple products within one OCT image, yielding a high number of total detections per test set, as described in the following section.

#### 4.3. Sensitivity: number of evaluations from defined test sets

A comparison of detection sensitivity highlights the excellent performance of the CNN approach. When evaluating highly scattering coating layers containing iron (II) oxide ( $\text{Fe}_2\text{O}_3$ ), the ellipse-fit could barely detect the valid interfaces. Therefore, the corresponding bars in Fig. 6 are very small.

#### 4.4. Challenging coating layer evaluation

Under the ellipse-fit based evaluation approach, the most challenging coating layer type is the one with a highly porous coating layer structure. This is especially the case for Glucotrol XL 10 mg tablets. Pores

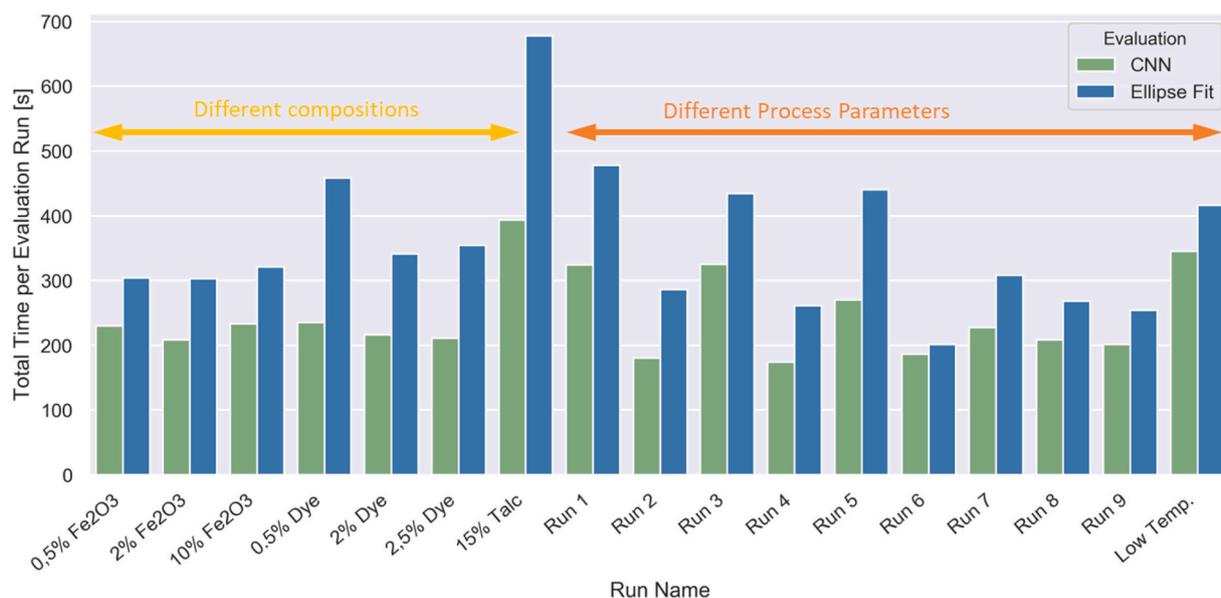


Fig. 5. Comparison of computational time required for CNN and for ellipse-fitting for a complete evaluation of selected in-line data sets of Thrombo ASS, covering a multitude of image appearances introduced by different process parameters and coating solution variations.

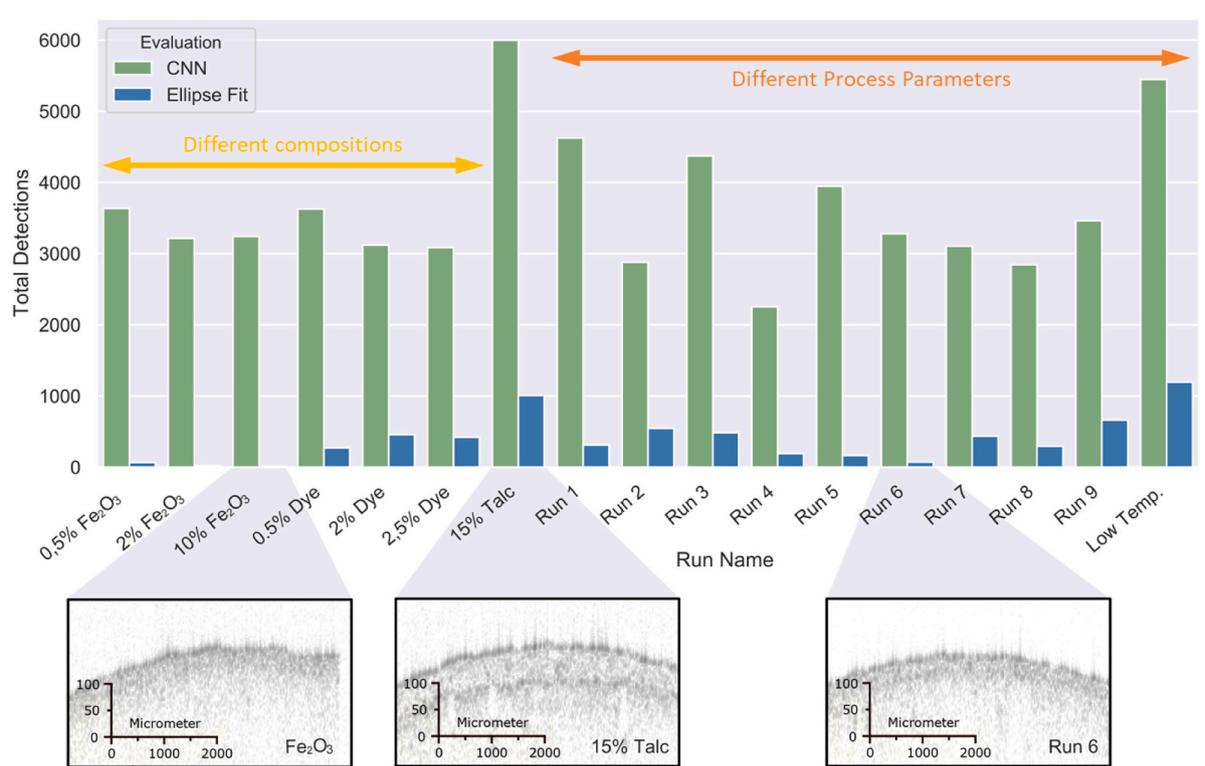
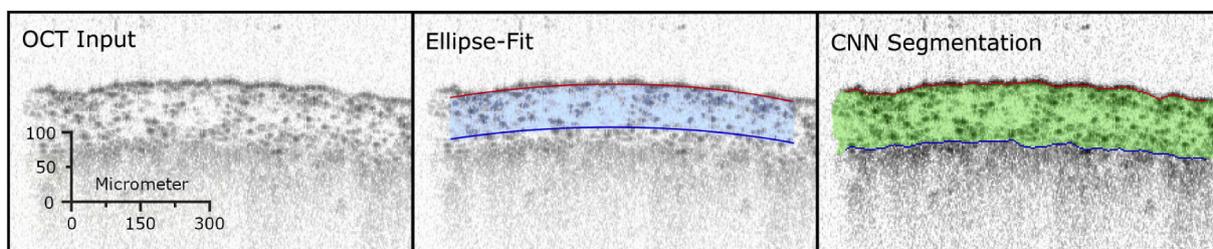


Fig. 6. Comparison of algorithmic sensitivity and robustness of CNN to that of ellipse-fitting for selected in-line data sets of Thrombo ASS, covering a multitude of image appearances introduced by different process parameters and coating solution variations.

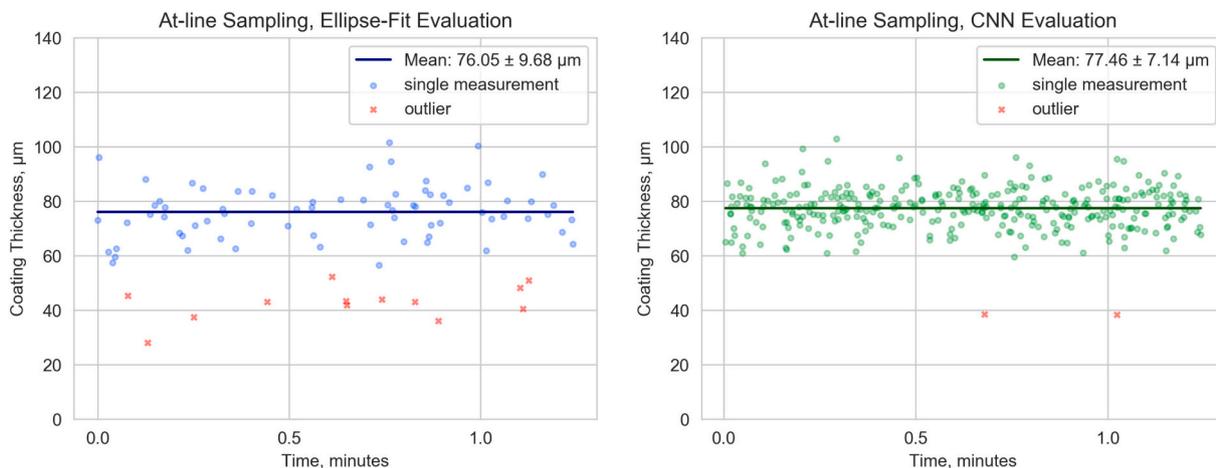
within the layer scatter the OCT light beam and produce a kind of visually “speckled” coating, which in most cases impedes a proper ellipse-fit evaluation by failing to recognize the coating-core interface. This is illustrated in Fig. 7, in which an ellipse-fit detection outlier is compared to the corresponding CNN-based segmentation output. In addition, Fig. 8 shows an at-line run comparison between both algorithms for the same product, with the ellipse-fit based evaluation and CNN-based algorithm yielding 86 and 329 detections, respectively, in 564 randomly sampled (empty and non-empty) full OCT images.

#### 4.5. In-line evaluation comparison

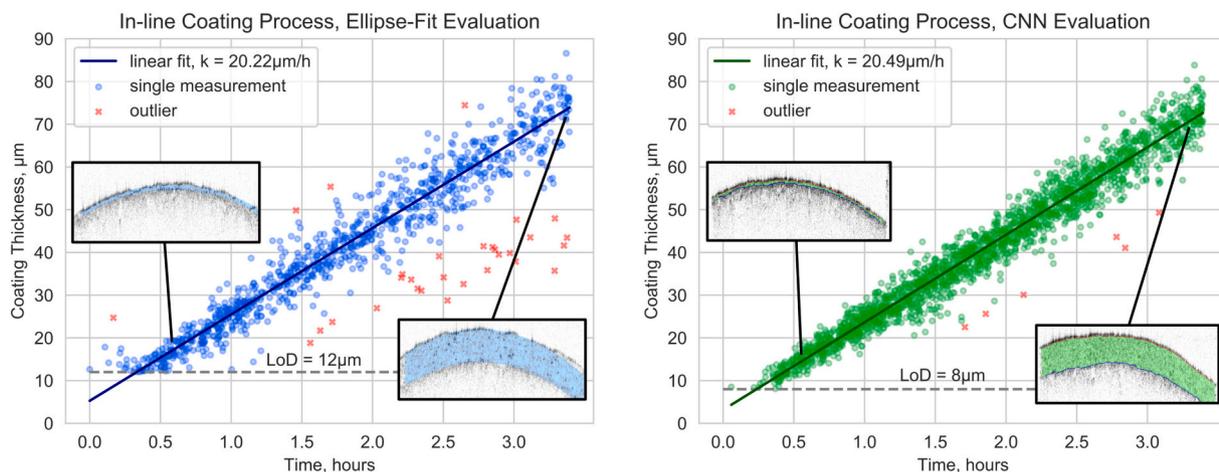
For the 15% Talcum run, the entire process evaluation of over more than 3 h coating time was re-processed based on recorded raw data (Fig. 9). The ellipse-fit reported 1007 coating layer detections, whereas CNN yielded 5998 detections from the same data set consisting of 7254 OCT images. The reason is that CNN is capable of evaluating multiple tablets in one OCT image while the static ellipse-fit can process only one tablet per OCT image. The CNN-based evaluation proved to be



**Fig. 7.** Coating layer detection on a Glucotrol XL 10 mg product example, with a highly porous structure showing an ellipse-fit outlier. The ellipse-fit approach fails to correctly detect the bottom coating-core interface. The CNN is capable of detecting the full layer area, determining correct interface paths with a much higher accuracy. The red and blue lines illustrate the top interface paths and the coating-core interface paths, respectively. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)



**Fig. 8.** At-line data evaluation of Glucotrol XL 10 mg tablets. Ellipse-fit evaluation (left) and CNN-based evaluation (right), including outliers (obtained via a review of all data by a human expert) and the batch mean thickness. The tablets were sampled using a commercially available Phyllon sampling device.



**Fig. 9.** In-line data evaluation for 15% Talcum run over more than 3 h of coating time. Ellipse-fit evaluation (left) and CNN-based evaluation (right), including outliers and linear fit. The results show individual thickness measurements over time and are based on a fixed dataset. The examples are shown for thin and thick layer detection.

sufficiently robust against changes in the coating layer thickness and layer appearance in the course of a coating run, as indicated by significantly fewer outliers (0.08%) compared to the ellipse-fit (2.88%) for the same test data.

The limit of detection for the ellipse-fit algorithm in terms of layer thickness is 12  $\mu\text{m}$ , which is determined by the optical resolution of the system in combination with the nature of static elliptical interface identification that can only evaluate in steps of full pixels. In contrast,

the CNN approach allows an evaluation of thinner layers (down to 8  $\mu\text{m}$ ), while still offering reliable detection. This is possible since interfaces are found by the CNN as areas and the evaluation is made on a subpixel level.

#### 4.6. Arbitrary shapes and multi-layers

Since the CNN-based layer segmentation proved to perform well for

single-layered tablets in terms of thickness and shape variety, multi-layered pellet evaluation was the logic next step. Fig. 10 shows an illustration of multi-layer segmentation with a pre-trained CNN model. The scanned objects are images of Detrol pellets acquired on a rotating Petri dish, with a static OCT sensor mounted above to prevent distortion of the OCT image due to random movement. Note that the CNN was not trained to learn the top layer (seal layer) since its thickness was represented by only a few pixels and was indistinguishable (for the human annotators) from the actual top interface.

The major advantage of CNN-based evaluation in this application is that the presented CNN is capable of following the real layer shape by simultaneously classifying each layer. Furthermore, multiple products in a single OCT image can be evaluated in parallel. In this case, the CNN model outputs three segmentation maps, one for each learned layer type.

## 5. Discussion

### 5.1. Comparison of algorithms

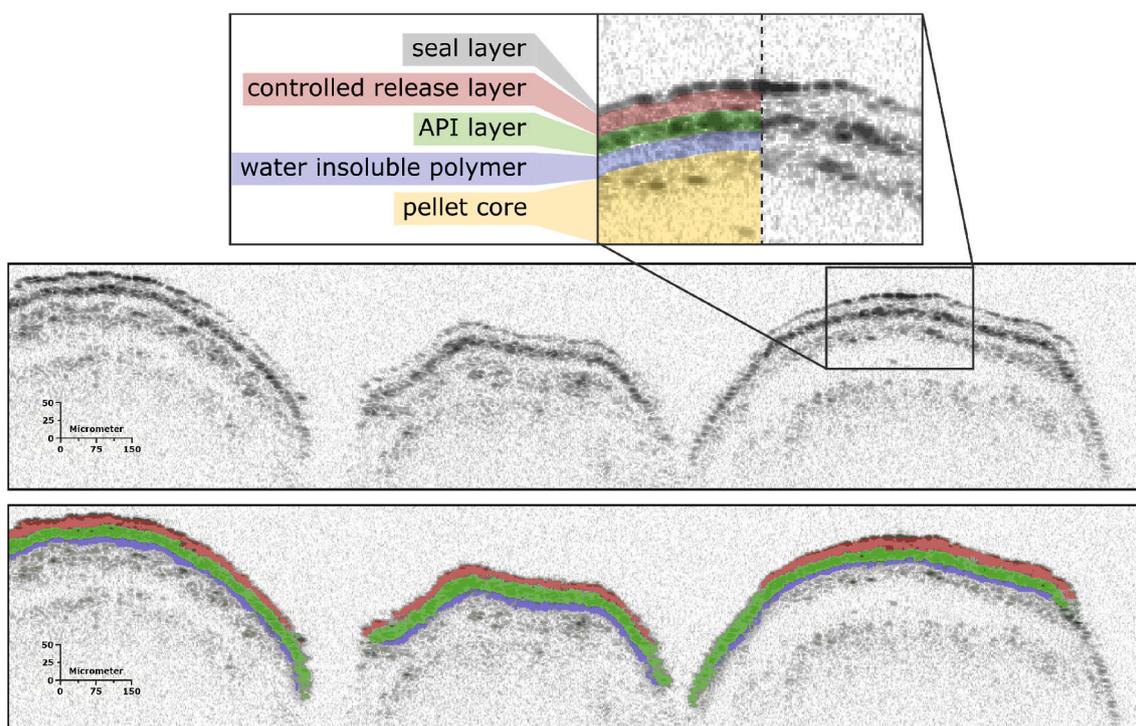
The presented CNN approach outperforms the ellipse-fit evaluation in terms of speed on the same hardware under all test scenarios and has enhanced robustness against a challenging textural appearance of porous coatings (e.g., Glucotrol XL tablets) and challenging shapes (e.g., the pellet examples described above). Furthermore, the total number of detections under the CNN approach for defined in-line data test sets is considerably higher than that under the ellipse-fit approach by a factor of 4.5 up to 324 for scattering layers, as shown in Fig. 6 for the various Thrombo ASS coating runs. This advantage is only partly due to the CNN's capability to segment multiple objects and parts of the trained objects at once. It holds true even when the objects appear mirrored in the image, while the ellipse-fitting can segment only one properly aligned object per OCT image. The major reason why the CNN outperforms the ellipse-fit is because it can recognize shapes and areas of interest, even with weak image contrast and high noise. Another notable

advantage of the CNN approach is that thinner coating layers (down to 8  $\mu\text{m}$ ) can be investigated on the same hardware since the CNN can evaluate areas and offers reliable sub-pixel information.

As the results summarized in Fig. 5 and especially Fig. 6 indicate, different compositions and varying process conditions introduce a high variability in the appearance of recorded OCT images. The developed CNN algorithm showed a high robustness even under these adverse circumstances, outperforming the ellipse-fit particularly in terms of highly scattering coatings containing inorganic pigments ( $\text{Fe}_2\text{O}_3$ ), in which case the ellipse-fit failed to evaluate the layers for 2% and 10% loading. Looking at the runs containing different amounts of dye, the difference in performance for both approaches was comparable to the runs without dye at the same process conditions, indicating neglectable influence of added dye on the resulting image quality.

Nevertheless, the CNN-based evaluation is still in a proof-of-concept state and requires enhanced post-processing, such as filtering false detections. The major challenge in training CNNs is that training data must be manually generated by providing a highly accurate annotation, which can be cumbersome work at the pixel level. However, as shown in Fig. 4, under the proposed image patch training approach only 30 training images suffice for achieving a dice score of 0.931 for the Thrombo ASS tablets.

Pellets are more challenging than tablets since they have a much thinner coating in general. Pellet cores can have arbitrary shapes, often causing the coated layer to have a non-spherical or non-elliptical curvature. As a consequence, the ellipse-fit approach reaches its limits when applied to such thin coatings, returning a comparatively high amount of false detections or even failing to segment a layer. However, these issues can be resolved using the proposed CNN-based segmentation (Fig. 10). The noticeably lower dice score of pellet layer detection can be explained by its dependency on the underlying areas. Since pellet layers are quite thin, discrepancies between the manually annotated ground-truth layer area and the CNN output segmentation area have a much higher impact on this value. Thus, the resulting dice score may be



**Fig. 10.** CNN-based multi-layer segmentation illustration for the Detrol pellets. (Top) OCT input image showing 3 pellet instances and an explanation of layers. (Bottom) Segmentation overlay of three CNN output maps, each using a confidence threshold of 0.5, denoting the controlled release layer in red, the API layer in green and the water-insoluble barrier layer in blue. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

considered to be a less significant metric in this special case.

Regarding multi-layered pellets, the deep learning approach was proven to be a key method for simultaneously evaluating multiple layer types in one product. This was expected since while learning a specific layer (e.g., the controlled release layer), the network can consolidate its prediction by using the surrounding image information for all other layers. This way, the other layers, interfaces or any additional ordered image structures support the CNN's learning. The ellipse-fit approach on the other hand often yields false detections due to mistaking layers that have a similar appearance.

Unlike the ellipse-fitting that relies on fixed parameters and thresholds, the CNN is capable of segmenting even faint interfaces with very weak contrast ratios. Nevertheless, the interfaces must still be visible to the human expert in order to train the network. Another aspect of the comparatively "rigid" architecture of static algorithms, such as ellipse-fitting, is that the image size and the framerate must match the prerequisites of the software. In contrast, the presented deep learning network remains flexible in terms of input and output image size and framerate and is only limited by the performance of available hardware. Even a continuous OCT image data stream could be fed in the CNN in order to evaluate every recorded A-scan. In other words, with the proposed CNN approach the performance directly scales with the computational power of the system it is running on.

### 5.2. Constrains and limitations of the CNN approach

One of the major drawbacks of CNN-based OCT image analysis is that a sufficient amount of training data must be created from scratch, especially in the case of coating layers with new types of visual appearance. This requires human experts to conduct the annotation of entire OCT images at the pixel level by marking the layer area as accurately as possible, which can be challenging if the interfaces are only weakly defined. Since the network's segmentation performance is primarily affected by the training data accuracy, this is a crucial point. In addition, the CNN models are in general highly sensitive to the input image noise, which can be a disadvantage. During live acquisition of OCT images, changes in the noise may occur, meaning that the segmentation may have unexpected behavior.

## 6. Summary and conclusions

Analyzing tablet or pellet coatings in OCT images during live acquisition is a challenging task not only due to the image quality constraints but also due to arbitrary object shapes and random object locations and movement. The latest ellipse-fit algorithm used to identify coating interfaces may have issues with that regard for non-elliptical products.

In this work, a proof of concept was developed for applying convolutional neural networks (CNNs) to coating layer segmentation and classification while still retaining the real-time characteristic of the overall evaluation algorithm. The proposed CNN-based coating layer segmentation has very high detection rates since it can simultaneously evaluate multiple products in one OCT image without additional computational effort. This also applies to multi-layered pellets, on which multiple coating layers were identified at once via a single CNN model. Another advantage of the novel approach is that scanned pharmaceuticals outside the ideal OCT image position (resulting in mirrored coating layers) can be learned and segmented without any additional algorithm adaptations. In contrast to the established static ellipse-fit algorithm, the CNN-based method requires fewer setting parameters, which makes it comparably straightforward for configuration and application. Validation of the results for defined test sets with challenging coatings (e.g., scattering and porous coatings) attested to the excellent performance and reliability of the CNN approach.

A limitation to a broad application of CNNs in the field of pharmaceutical OCT image evaluation is that no readily available training data

for pharmaceutical OCT image segmentation is currently available. Training data must be generated manually by human experts and for every application. As a consequence, annotation errors or bias can influence the final CNN-based segmentation accuracy, which may also affect the measurement validity.

Future work will include examining the overall detection reliability of and further improvements to the CNNs' training and robustness. Learning additional product properties, such as pore presence, could enable a calculation of inner surfaces of the coating and improve the overall layer homogeneity calculation. Another aspect with regard to CNNs would be to learn and correct the individual light beam inclination of live OCT data, which may be derivable from the top interface reflection appearance. This could make possible a live compensation for the refraction effects based on the unseen surface angle and a correction of the final layer thickness measurement.

### CRedit authorship contribution statement

Conceptualization: Matthias Wolfgang, Michael Weißensteiner  
 Data curation: Matthias Wolfgang, Michael Weißensteiner  
 Formal analysis: Matthias Wolfgang, Michael Weißensteiner, Wen-Kai Hsiao  
 Funding acquisition: Johannes G. Khinast  
 Investigation: Matthias Wolfgang, Michael Weißensteiner  
 Methodology: Matthias Wolfgang, Michael Weißensteiner  
 Project administration: Phillip Clarke  
 Resources: Phillip Clarke, Johannes G. Khinast  
 Software: Michael Weißensteiner  
 Supervision: Wen-Kai Hsiao, Johannes G. Khinast  
 Validation: Matthias Wolfgang  
 Visualization: Matthias Wolfgang, Michael Weißensteiner  
 Writing - original draft: Matthias Wolfgang, Michael Weißensteiner  
 Writing - review & editing: Matthias Wolfgang, Michael Weißensteiner, Phillip Clarke, Wen-Kai Hsiao, Johannes G. Khinast

### Declaration of Competing Interest

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

### Acknowledgements

This work has been funded by the Austrian COMET Program under the auspices of the Austrian Federal Ministry of Transport, Innovation and Technology (bmvit), the Austrian Federal Ministry of Economy, Family and Youth (bmfj) and by the State of Styria (Styrian Funding Agency SFG). COMET is managed by the Austrian Research Promotion Agency FFG.

Moreover, the authors would like to thank Simon Pöllitsch for his valuable contribution during algorithm development and optimization, as well as Anna Peter and Vanessa Herndler for their help and dedication during OCT image annotation, testing and validation of the different versions of algorithms. Supply of pellet and tablet samples for our study by Pfizer Inc. is also acknowledged.

### References

- Chollet, F., et al., 2015. Keras [WWW Document]. URL. <https://github.com/fchollet/keras>.
- de Boer, J.F., Leitgeb, R., Wojtkowski, M., 2017. Twenty-five years of optical coherence tomography: the paradigm shift in sensitivity and speed provided by Fourier domain OCT. *Biomed. Opt. Expr.* 8, 3248. <https://doi.org/10.1364/BOE.8.003248>.
- del Amor, R., Morales, S., Colomer, A., Mogensen, M., Jensen, M., Israelsen, N.M., Bang, O., Naranjo, V., 2020. Automatic segmentation of epidermis and hair follicles in optical coherence tomography images of normal skin by convolutional neural networks. *Front. Med.* 7, 1–11. <https://doi.org/10.3389/fmed.2020.00220>.

- Evonik Industries AG, 2011. EUDRAGIT® L 30 D-55 Enteric Coating with Talc as Anti-Tacking Agent 2.
- He, K., Zhang, X., Ren, S., Sun, J., 2015. Delving deep into rectifiers: Surpassing human-level performance on imagenet classification. In: Proc. IEEE Int. Conf. Comput. Vis. 2015 Inter, pp. 1026–1034. <https://doi.org/10.1109/ICCV.2015.123>.
- Kepp, T., Droigk, C., Casper, M., Evers, M., Hüttmann, G., Salma, N., Manstein, D., Heinrich, M.P., Handels, H., 2019. Segmentation of mouse skin layers in optical coherence tomography image data using deep convolutional neural networks. *Biomed. Opt. Expr.* 10, 3484. <https://doi.org/10.1364/boe.10.003484>.
- Kingma, D.P., Ba, J.L., 2015. Adam: A method for stochastic optimization. In: 3rd Int. Conf. Learn. Represent. ICLR 2015 - Conf. Track Proc, pp. 1–15.
- Lin, H., Dong, Y., Shen, Y., Zeitler, J.A., 2015. Quantifying pharmaceutical film coating with optical coherence tomography and terahertz pulsed imaging: an evaluation. *J. Pharm. Sci.* 104, 3377–3385. <https://doi.org/10.1002/jps.24535>.
- Lin, H., Dong, Y., Markl, D., Williams, B.M., Zheng, Y., Shen, Y., Zeitler, J.A., 2016. Measurement of the intertablet coating uniformity of a pharmaceutical pan coating process with combined terahertz and optical coherence tomography in-line sensing. *J. Pharm. Sci.* 106, 1075–1084. <https://doi.org/10.1016/j.xphs.2016.12.012>.
- Lin, H., Dong, Y., Markl, D., Zhang, Z., Shen, Y., Zeitler, J.A., 2017. Pharmaceutical film coating catalog for spectral domain optical coherence tomography. *J. Pharm. Sci.* 106, 3171–3176. <https://doi.org/10.1016/j.xphs.2017.05.032>.
- Lin, H., Zhang, Z., Markl, D., Zeitler, J., Shen, Y., 2018. A review of the applications of OCT for analysing pharmaceutical film coatings. *Appl. Sci.* 8, 2700. <https://doi.org/10.3390/app8122700>.
- Long, J., Shelhamer, E., Darrell, T., 2015. Fully convolutional networks for semantic segmentation. In: 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 3431–3440. <https://doi.org/10.1109/CVPR.2015.7298965>.
- Markl, D., Hanneschläger, G., Sacher, S., Leitner, M., Khinast, J.G., 2014. Optical coherence tomography as a novel tool for in-line monitoring of a pharmaceutical film-coating process. *Eur. J. Pharm. Sci.* 55, 58–67. <https://doi.org/10.1016/j.ejps.2014.01.011>.
- Markl, D., Hanneschläger, G., Sacher, S., Leitner, M., Buchsbaum, A., Pescod, R., Baele, T., Khinast, J.G., 2015a. In-line monitoring of a pharmaceutical pan coating process by optical coherence tomography. *J. Pharm. Sci.* 104, 2531–2540. <https://doi.org/10.1002/jps.24531>.
- Markl, D., Hanneschläger, G., Sacher, S., Leitner, M., Khinast, J.G., Buchsbaum, A., 2015b. Automated pharmaceutical tablet coating layer evaluation of optical coherence tomography images. *Meas. Sci. Technol.* 26, 1–12. <https://doi.org/10.1088/0957-0233/26/3/035701>.
- Markl, D., Wahl, P., Pichler, H., Sacher, S., Khinast, J.G., 2018. Characterization of the coating and tablet core roughness by means of 3D optical coherence tomography. *Int. J. Pharm.* 536, 459–466. <https://doi.org/10.1016/j.ijpharm.2017.12.023>.
- Neff, T., Payer, C., Stern, D., Urschler, M., 2017. Generative Adversarial Network based Synthesis for Supervised Medical Image Segmentation. *Proc. OAGM&ARW Jt. Work. 2017 Vision, Autom. Robot*, pp. 140–145. <https://doi.org/10.3217/978-3-85125-524-9-30>.
- Nemeth, A., Hanneschläger, G., Leiss-Holzinger, E., Wiesauer, K., Leitner, M., 2013. Optical Coherence Tomography – Applications in Non-Destructive Testing and Evaluation. *Intech*, pp. 163–185. <https://doi.org/10.5772/53960>.
- Radford, A., Metz, L., Chintala, S., 2016. Unsupervised representation learning with deep convolutional generative adversarial networks. In: 4th Int. Conf. Learn. Represent. ICLR 2016 - Conf. Track Proc, pp. 1–16.
- Ronneberger, O., Fischer, P., Brox, T., 2015. U-net: Convolutional networks for biomedical image segmentation. In: *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, 9351, pp. 234–241. [https://doi.org/10.1007/978-3-319-24574-4\\_28](https://doi.org/10.1007/978-3-319-24574-4_28).
- Rouba, B.J., Karaszkievicz, P., Tymiąska-Widmer, L., Iwanicka, M., Góra, M., Kwiatkowska, E., Targowski, P., 2008. Optical Coherence Tomography for Non-Destructive Investigations of Structure of Objects of Art. 9th Int. Conf. NDT Art 25–30.
- Roy, A.G., Conjeti, S., Karri, S.P.K., Sheet, D., Katouzian, A., Wachinger, C., Navab, N., 2017. ReLayNet: retinal layer and fluid segmentation of macular optical coherence tomography using fully convolutional networks. *Biomed. Opt. Expr.* 8, 3627. <https://doi.org/10.1364/boe.8.003627>.
- Rueden, C.T., Schindelin, J., Hiner, M.C., DeZonia, B.E., Walter, A.E., Arena, E.T., Eliceiri, K.W., 2017. ImageJ2: ImageJ for the next generation of scientific image data. *BMC Bioinformatics* 18, 1–26. <https://doi.org/10.1186/s12859-017-1934-z>.
- Sacher, S., Wahl, P., Weißensteiner, M., Wolfgang, M., Pokhilchuk, Y., Looser, B., Thies, J., Raffa, A., Khinast, J.G., 2019. Shedding light on coatings: real-time monitoring of coating quality at industrial scale. *Int. J. Pharm.* 566, 57–66. <https://doi.org/10.1016/j.ijpharm.2019.05.048>.
- Wahl, P.R., Peter, A., Wolfgang, M., Khinast, J.G., 2019. How to measure coating thickness of tablets: method comparison of optical coherence tomography, near-infrared spectroscopy and weight-, height- and diameter gain. *Eur. J. Pharm. Biopharm.* <https://doi.org/10.1016/j.ejpb.2019.06.021>.
- Wise, Donald L., 2000. *Handbook of Pharmaceutical Controlled Release Technology*, 1st ed. CRC Press.
- Wolfgang, M., Peter, A., Wahl, P., Markl, D., Zeitler, J.A., Khinast, J.G., 2019. At-line validation of optical coherence tomography as in-line/at-line coating thickness measurement method. *Int. J. Pharm.* 572, 118766. <https://doi.org/10.1016/j.ijpharm.2019.118766>.
- Zeitler, J.A., Gladden, L.F., 2009. In-vitro tomography and non-destructive imaging at depth of pharmaceutical solid dosage forms. *Eur. J. Pharm. Biopharm.* 71, 2–22. <https://doi.org/10.1016/j.ejpb.2008.08.012>.
- Zettl, M., 2017. *Investigation of a Continuous Tablet Coater Via Optical Coherence Tomography*. Technical University Graz.
- Zhang, Y., Qiu, Z., Yao, T., Liu, D., Mei, T., 2018. Fully convolutional adaptation networks for semantic segmentation. *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.* 6810–6818. <https://doi.org/10.1109/CVPR.2018.00712>.
- Zhong, S., Shen, Y.C., Ho, L., May, R.K., Zeitler, J.A., Evans, M., Taday, P.F., Pepper, M., Rades, T., Gordon, K.C., Miller, R., Kleinebudde, P., 2010. Non-destructive quantification of pharmaceutical tablet coatings using terahertz pulsed imaging and optical coherence tomography. *Opt. Lasers Eng.* 49, 361–365. <https://doi.org/10.1016/j.optlaseng.2010.11.003>.