BRNO UNIVERSITY OF TECHNOLOGY
VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ

FACULTY OF MECHANICAL ENGINEERING
FAKULTA STROJNÍHO INŽENÝRSTVÍ

INSTITUTE OF MACHINE AND INDUSTRIAL DESIGN
ÚSTAV KONSTRUOVÁNÍ

THE USE OF DEEP NEURAL NETWORKS FOR THE EVALUATION OF METALLOGRAPHIC CROSS-SECTIONS
VYUŽITÍ HLUBOKÝCH NEURONOVÝCH SÍTÍ PRO VYHODNOCENÍ METALOGRAFICKÝCH VÝBRUSŮ

MASTER'S THESIS
DIPLOMOVÁ PRÁCE

AUTHOR
AUTOR PRÁCE
Bc. Adam Semančík

SUPERVISOR
VEDOUCÍ PRÁCE
Ing. Jakub Hurník, Ph.D.

BRNO 2024

Assignment Master's Thesis

Institut:	Institute of Machine and Industrial Design
Student:	**Bc. Adam Semančík**
Degree programm:	Mechanical Engineering Design
Branch:	no specialisation
Supervisor:	**Ing. Jakub Hurník, Ph.D.**
Academic year:	2023/24

As provided for by the Act No. 111/98 Coll. on higher education institutions and the BUT Study and Examination Regulations, the director of the Institute hereby assigns the following topic of Master's Thesis:

The use of deep neural networks for the evaluation of metallographic cross–sections

Brief Description:

Indicators such as porosity, micro–cracks and other important characteristics are used to evaluate the quality of metal 3D printing using SLM technology. These characteristics are currently evaluated from metallographic cross–sections using simple image analysis methods such as thresholding. Moreover, the parameters of this method are determined only by user estimation or automatically based on simple histogram characteristics. This leads to an inaccurate result, and the method is also highly sensitive to possible imperfections in the image.

Thesis type: development
Thesis result: software (R), publication (J, D)
Project: specific research

Master's Thesis goals:

The aim of this work is to propose a method for evaluating the characteristics of metallographic cross–sections based on advanced image processing methods, e.g. using neural networks.
Sub–objectives of the thesis:
– research,
– selection of potentially suitable methods,
– implementation of methods,
– experimental validation.

Required outputs: thesis, digital data.
Work extent: approx. 72 000 characters (40 – 50 pages of text without figures).
The timetable, the structure of the work and the thesis template are obligatory:
http://www.ustavkonstruovani.cz/texty/bakalarske–studium–ukonceni/

Recommended bibliography:

CHRIS, Solomon a Toby BRECKON. Fundamentals of Digital Image Processing: a practical approach with examples in matlab. Oxford: Wiley-Blackwell, 2011. ISBN 978-047-0844-731.

GE, M., F. SU, Z. ZHAO a D. SU. Deep learning analysis on microscopic imaging in materials science. Materials Today Nano [online]. 2020, 11. ISSN 25888420. In: doi:10.1016/j.mtnano.2020.100087.

DONG, Chao, Chen Change LOY, Kaiming HE a Xiaoou TANG. Image Super-Resolution Using Deep Convolutional Networks. IEEE Transactions on Pattern Analysis and Machine Intelligence [online]. 2016, 38(2), 295–307. ISSN 01628828. In: doi:10.1109/TPAMI.2015.2439281

Deadline for submission Master's Thesis is given by the Schedule of the Academic year 2023/24

In Brno,

L. S.

_____   _____
prof. Ing. Martin Hartl, Ph.D.   doc. Ing. Jiří Hlinka, Ph.D.
Director of the Institute   FME dean

ABSTRACT

This thesis explores the application of deep neural networks to improve the evaluation of metallographic cross-sections in materials produced through powder bed fusion. It focuses on two advanced image processing techniques: semantic segmentation and image super-resolution. A U-Net architecture was used for semantic segmentation to classify defects such as lack of fusion porosity and gas porosity. Additionally, an SRGAN (Super-Resolution Generative Adversarial Network) model was utilized to upscale image resolution, potentially enhancing segmentation accuracy. The research assesses whether a model trained on AlSi10Mg can generalize to Cu99 and Ti6Al4V and evaluates the influence of super-resolution on segmentation performance. Results showed that while the segmentation model performed well on AlSi10Mg, generalization to other materials required more diverse training data. Due to computational limitations, the combined effect of super-resolution and segmentation remains inconclusive, suggesting further research with enhanced computational resources.

KEYWORDS

deep neural networks, semantic segmentation, super-resolution, metallographic cross-sections, powder bed fusion, additive manufacturing, U-net, SRGAN, image processing, machine learning

ABSTRAKT

Táto diplomová práca skúma aplikáciu hlbokých neurónových sietí pre vylepšenie hodnotenia metalografických výbrusov pre materiály vyrobené pomocou aditívnej výroby. Zameriava sa na dve pokročilé techniky spracovania obrazu: sémantickú segmentáciu a super-rozlíšenie obrazu. Na sémantickú segmentáciu bola použitá architektúra U-Net pre klasifikáciu defektov, ako sú dva typy pórov. Okrem toho bol použitý model SRGAN (Super-Resolution Generative Adversarial Network) pre zvýšenie rozlíšenia obrazu, čo potenciálne zlepšuje presnosť segmentácie. Výskum hodnotí, či model trénovaný na AlSi10Mg môže dostatočne dobre vyhodnocovať materiály Cu99 a Ti6Al4V. Zároveň hodnotí vplyv super-rozlíšenia na výkonnosť segmentácie. Výsledky ukázali, že zatiaľ čo model segmentácie dosahoval dobré výsledky na AlSi10Mg, generalizácia na iné materiály vyžaduje diverzifikovanejšie tréningové dáta. V dôsledku výpočtových obmedzení zostáva kombinovaný efekt super-rozlíšenia a segmentácie nejednoznačný, čo naznačuje potrebu ďalšieho výskumu s výkonnejšími výpočtovými zdrojmi.

KĽÚČOVÉ SLOVÁ

hlboké neurónové siete, sémantická segmentácia, super-rozlíšenie, metalografické výbrusy, aditívna výroba, U-net, SRGAN, spracovanie obrazu, strojové učenie

BIBLIOGRAPHIC REFERENCE

SEMANČÍK, Adam. *The use of deep neural networks for the evaluation of metallographic cross-sections*. Brno, 2024. Available also at: https://www.vut.cz/studenti/zav-prace/detail/154153. Diploma thesis. Brno University of Technology, Faculty of Mechanical Engineering, Institute of Machine and Industrial Design. Thesis supervisor Jakub Hurník.

AUTHOR'S DECLARATION ON THE ORIGIN OF THE WORK

CONTENT

1 INTRODUCTION

Computer vision is a topic that studies how to obtain visual information of world around us to help computers understand it better. With rise of deep learning, since 2012, these methods based on neural networks managed to significantly improve the state-of-art in object detection, object recognition and segmentation that can nowadays outperform human experts in their fields [1].

For evaluation and optimization of process parameters in metal additive manufacturing, different parameters are used. These are nowadays still obtained from metallographic cross-sections by simple image processing methods like threshold-based segmentation or edge-based segmentation. These methods are often sensitive to noise, imperfections, or more complicated textures of cross-sections.

Applications in metallography of metal additive manufacturing leads to the application of neural networks and deep learning, for the goal of improving materials and process parameters [2].

First, I need to point out that there are three major approaches to deep learning.

- **Deep supervised learning** - This method works with data labelled by humans. That means that we 'teach the model' iteratively by providing input and wanted output, so it can than predict outcome of out-of-sample data by itself. Generally, this approach is simpler and more transparent in terms of coding and understanding its behavior.
- **Deep unsupervised learning** - This technique can be used to make model learn by itself. That means that models can automatically find connections between input and output data.
- **Deep semi-supervised learning** - This approach is both mentioned techniques together. At the beginning the model is trained on labelled data and then it can learn by itself from out-of-sample data. For example, the most popular application of this method is text document classifier [3].

Semantic segmentation is used for pixel level image classification. Output of semantic segmentation neural network is classification map with information about which pixel belongs to which instance [4]. Luengo et. al. has recently tested deep learning approaches of supervised, unsupervised, and semi-supervised learning for this approach of semantic segmentation of metallographic images and shown, that deep supervised learning outperforms unsupervised and semi-supervised techniques [2]. That is the reason why I am focusing only on supervised learning methods in this thesis.

Image super-resolution is a deep learning-based method for upscaling image resolution to enhance image quality and details. Super-resolution has shown great success in the past years and this technique is suitable for use in metallographic analysis with the goal of extracting more details from image samples, by taking images with lower resolution. Generally, the main aim of image super-resolution is to enhance image resolution and image quality to human visual system. That's why there are many studies centered around this purpose. However, previous studies have shown that this approach can be helpful for many different computer vision tasks. Especially GAN (Generative Adversarial Network) based super-resolution techniques are aiming to recover sharp edges of image instances. That is the reason why I am going to focus on GAN based, supervised super-resolution techniques in this thesis [5].

The goal of this thesis is to use deep learning based neural network algorithms for evaluation of metallographic cross sections. This approach should bring more information from the image of cross-sections and provide better understanding of what is happening inside of material during 3D printing with powder bed fusion process.

2 STATE OF THE ART

2.1 Systematic review methods

The aim of this review was to provide the most reliable and the most suitable publications for further development of diploma theses. Review should bring new knowledge about applications of Super-resolution and Semantic Segmentation in the field of metallographic image processing and evaluation. For this purpose, I have used principles of preferred reporting items for systematic reviews and Meta-Analyses (PRISMA). The step-by-step process of this approach is shown in PRISMA diagram (Fig. 2-1).
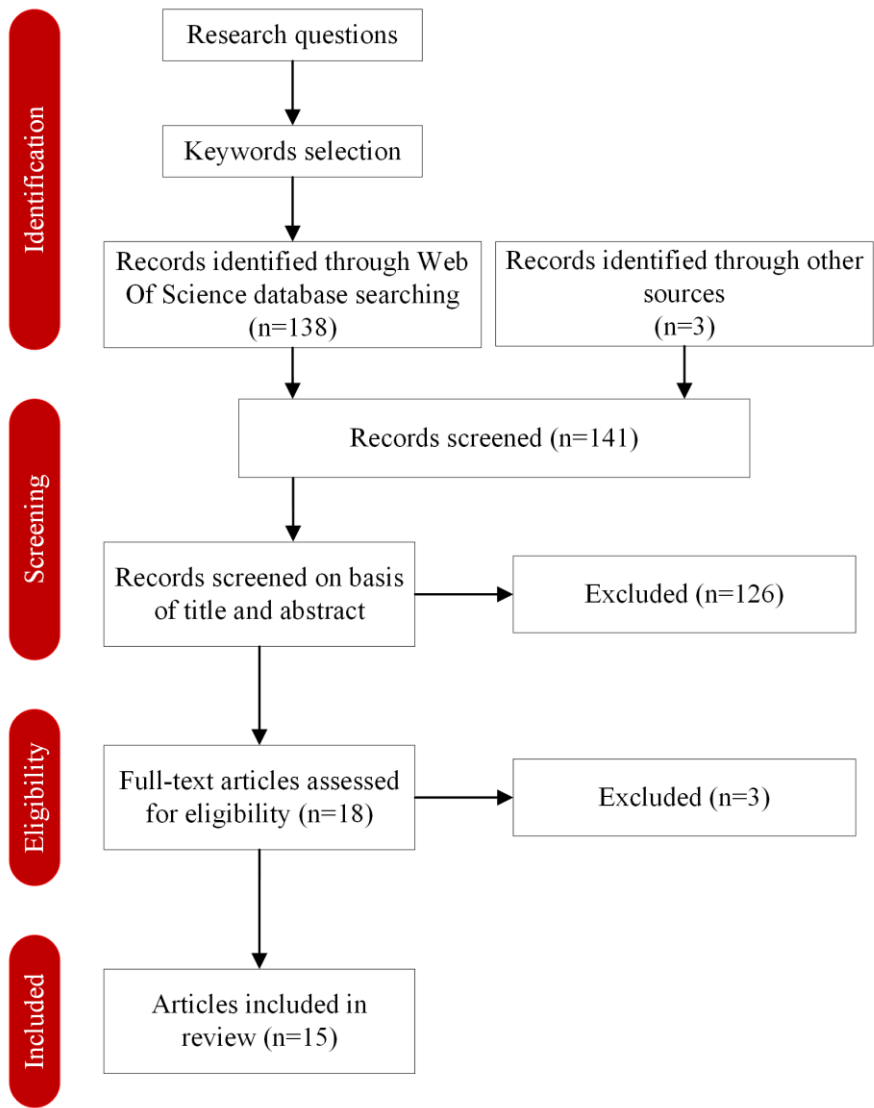
Fig. 2-1   PRISMA diagram.

I have created research questions (Q) shown in Tab. 2-1 to guide the review process. Q1 was created to find most recent articles about research in metallography of additive manufacturing to find out if anyone recently used deep learning or any kind of more sophisticated method to evaluate additive manufactured materials. Q2 discusses recent applications of deep learning with Semantic Segmentation and Super-Resolution in metallography. Q3 surveys whole area of neural networks for Semantic Segmentation. Similarly, Q4 discusses the area of image Super-Resolution.

Tab. 2-1   Research questions for systematic review.

Id	Issue
Q1	What methods are used for evaluating metallographic cross sections made by additive manufacturing?
Q2	Are there any applications of Semantic Segmentation or Super-Resolution in metallography?
Q3	What types of neural network architectures are used for Semantic Segmentation?
Q4	What types of neural network architectures are used for image Super Resolution?

After I created research questions, keywords were generated to each question for database searching. Then I defined the search string to each research question using basic Boolean operators, keywords, and their synonyms. As a searching database, Web of Science was chosen and I have focused on review articles, journal articles and conference proceedings. The idea was to find many results with smaller search string and then iteratively improve and trim results to smaller numbers by adding more keywords and Boolean operators. Final research strings with their number of results are shown in Tab. 2-2.

Tab. 2-2   Search strings for systematic review.

Id	Issue	Results
Q1	("metal printing" OR "metal additive manufacturing" OR "metal 3D printing") AND ("evaluation" OR "analysis") AND ("metallurgy" OR "metallography")	16
Q2	("metallography" OR "metallurgy") AND ("neural networks" OR "deep learning" OR "machine learning") AND applications AND ("picture" OR "image" OR "photography")	18
Q3	("neural networks" OR "deep learning" OR "machine learning") AND ("semantic segmentation" OR "instance segmentation") AND "network* architecture*" AND pixel	63
Q4	("neural networks" OR "deep learning" OR "machine learning") AND "single image* super-resolution" AND "network* architecture*"	41

After I have identified records to all research questions, they were screened by reading their titles and abstracts. I have mainly focused on finding results that are relevant in finding answers to research questions and I have selected 18 articles.

Then I proceeded to full-text articles. At first, I provided bibliometric information about each paper, to have some first idea about authors and journals before I started reading articles. I have provided impact factor and quartile of journal, where the article was published same as number of citations of article. These bibliometric data is shown in Tab. 2-3. Overall, after reading full-text articles, 15 publications were selected for this systematic review.

Tab. 2-3 Selected publications sorted by number of citations.

Publication	Type	IF	Quart.	Year	Cit.
Photo-realistic single image super-resolution using a generative adversarial network [6]	Conf.			2017	3146
ESRGAN: Enhanced super-resolution generative adversarial networks [7]	Conf.			2019	252
Review of deep learning: concepts, CNN architectures, challenges, applications, future directions [3]	Review article	11.1	Q1	2021	115
High throughput quantitative metallography for complex microstructures using deep learning: A case study in ultrahigh carbon steel [8]	Journal article	4.1	Q2	2019	66
Metallurgy, mechanistic models and machine learning in metal printing [9]	Review article	66.3	Q1	2021	61
A Metallographic Review of 3D Printing/Additive Manufacturing of Metal and Alloy Products and Components [10]	Review article	1.3	Q2	2018	50
Deep Learning for Semantic Segmentation of Defects in Advanced STEM Images of Steels [11]	Journal article	4.4	Q1	2019	38
Resolution enhancement in scanning electron microscopy using deep learning [12]	Journal article	4.4	Q1	2019	35
A review of deep learning methods for semantic segmentation of remote sensing imagery [4]	Review article	7	Q1	2021	34
Deep Learning-Based Weld Contour and Defect Detection from Micrographs of Laser Beam Welded Semi-Finished Products [13]	Journal article	2.7	Q2	2022	7
A generic high-throughput microstructure classification and quantification method for regular SEM images of complex steel microstructures combining EBSD labeling and deep learning [14]	Journal article	8.1	Q1	2021	5
Real-world single image super-resolution: A brief review [5]	Review article	13	Q1	2022	3
A deep learning method for extensible microstructural quantification of DP steel enhanced by physical metallurgy-guided data augmentation [15]	Journal article	4.3	Q1	2021	0
A tutorial on the segmentation of metallographic images: Taxonomy, new MetalDAM dataset, deep learning-based ensemble model, experimental analysis and challenges [2]	Journal article	13	Q1	2022	0
Super-resolution reconstruction, recognition, and evaluation of laser confocal images of hyperaccumulator Solanum nigrum endocytosis vesicles based on deep learning: Comparative study of SRGAN and SRResNet [16]	Journal article	5,6	Q1	2023	0

2.2 Critical systematic review

2.2.1 Metallography and machine learning in additive manufacturing

Murr L reviewed recent progress in metallography of additive manufactured materials [10]. DebRoy T et. al. recently pointed out that machine learning has been already widely used in additive manufacturing for different tasks by the goal of improving process parameters, tailoring microstructures and properties or defect reduction [9].

2.2.2 Semantic segmentation in metallography

In past years there have been some applications of deep convolutional neural network architectures for Semantic Segmentation in metallography that I need to point out here.

DeCost et. al. applied PixelNet [17] segmentation architecture, aiming to teach this model to find phases of ultrahigh carbon steel. This architecture can learn representation of each pixel of image fast, so it is suitable for smaller datasets. At the same time, they were experimenting with popular U-Net [18] architecture but they found out that it slightly underperformed PixelNet architecture. They were comparing two kinds of training loss functions. Standard cross-entropy loss and focal loss, which is designed for unbalanced datasets [8].

For training they used 24 manually annotated grayscale micrographs from their own open-sourced UHCS (Ultra High Carbon Steel) dataset [19]. For image labelling they used Medical Imaging Interaction Toolkit (MITK) [20] which is open-source image and voxel mesh annotating tool, mainly developed for medical use. Micrographs had 645x484 pixels and we can see their examples on Fig. 2-2. To extend the size of the dataset, they used common data augmentation techniques. They used random rotation in range $[0, 2\pi)$, vertical and horizontal flip, scaling in range $[1,2]$, and a $\pm5\%$ random intensity shift [8].
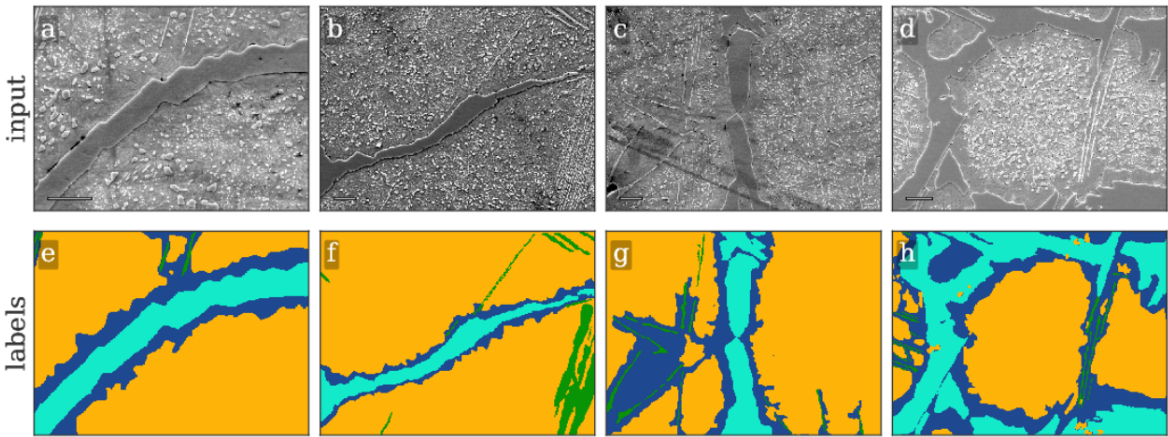

Fig. 2-2   Examples of UHCS dataset (a-d) micrographs, (e-h) labeled data [8].

They classified 4 different instances of ultrahigh carbon steel (matrix, network, spheroidite, widmanstatten). To evaluate performance, they used pixel accuracy (PA) region intersection over union (IoU), precision and recall. Overall, they managed to obtain an average PA of 91,1% with their best approach with cross-entropy loss function [8].

Roberts G et. al. wasn't using already created architecture. They created own model called DefectSegNet for Semantic Segmentation of three common metallographic defects in structural alloys: dislocation lines, voids, and precipitates. DefectSegNet draws inspiration from the well-known U-Net architecture. It incorporates skip connections both within and across blocks, which is the main difference from the traditional U-Net architecture. This modification has been found advantageous for their specific use case. [11].

To generate data for further training they used diffraction contrast imaging with scanning transmission electron microscope, and they used 250 000× magnification and 2048×2048 pixels image size. They manually labelled 2 micrographs of this size. On Fig. 2-3 is shown how they augmented dataset by splitting each of two micrographs to 3 training subsets (each 1024×1024 pixels), one development (validation) set and one test set (each 1024×512 pixels). Then they rotated each image of training subsets by 0.5π, π and 1.5π and finally flipped every rotated image horizontally. In total they generated 48 of 1024×1024 labelled images in training subset [11].
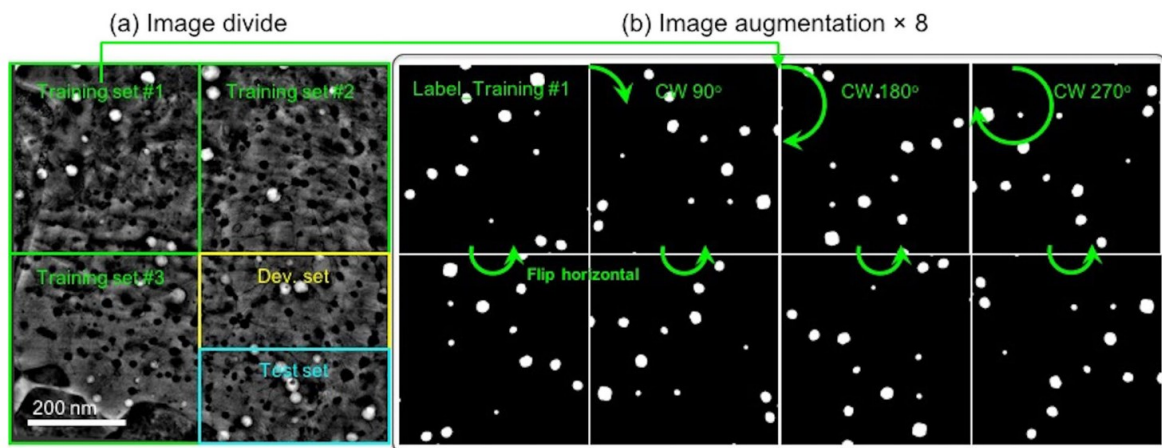

Fig. 2-3   Indication of (a) division and (b) augmentation of a pre-processed images [11].

They have used PA, precision, recall and IoU as evaluation metrices. They have managed to obtain over-all average of 94,6% PA and 61,79 IoU [11].

Shen C et. al. used mentioned U-Net [18] model, due to its great performance on small datasets, to find correlation between regular scanning electron microscope image and image of same frame taken by electron backscattered diffraction (EBSD), which is method of providing information about crystal orientation, structure, strain, or phase of structural materials [14].

Available input images came from two engineering steels. Dual-phase steel (austenite and martensite) and quenching and partitioning steel (ferrite, martensite and retained austenite. From these two steels they created two datasets. In the first dual-phase steel dataset, they created 248 images of 128×128 pixels. For quenching and partitioning steel dataset, they created 726 images of 128×128 pixels. Examples of images from dataset are shown on Fig. 2-4. For these two datasets they used similar augmentation method as have been mentioned (flipping and random cropping) and they increased amount of data in datasets to 1914 dual-phase (DP) steel paired pictures and 6048 quenching and partitioning (Q&P) steel paired pictures [14].
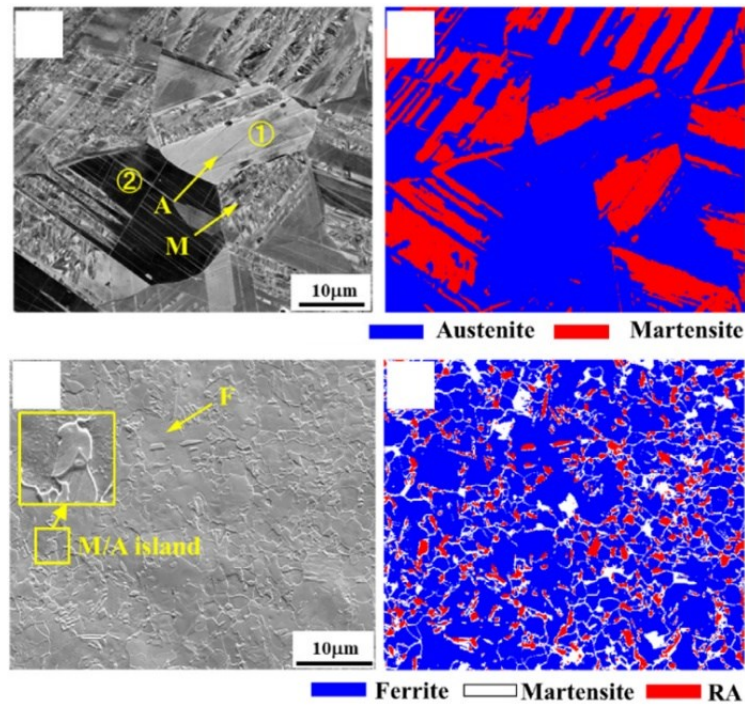

Fig. 2-4   (a) dual-phase steel, (b) electron backscattered diffraction image of dual-phase steel, (c) quenching and partitioning steel, (d) electron backscattered diffraction image of quenching and partitioning steel [14].

For evaluation, they used PA and IoU, achieving averages of 85.4% and 75.5%, respectively.

Same authors with same research group Shen C et. al. in their second work were using same U-Net [18] architecture to quantify grain size and fractions of martensite and ferrite in dual-phase steels.

They used field emission scanning electron microscopy with 1000× magnification to obtain micrographs. This time they focused on dual-phase steel with martensite and ferrite. In the dataset they had 100 subimages in total for model training subset and 40 subimages for testing subset, and all of them had 128×128 pixels. As before, they used similar dataset augmentation methods (random rotation, transformation, brightness adjustment). In total, they generated 2000 subimages for the training subset. Examples are shown on Fig. 2-5 [15].
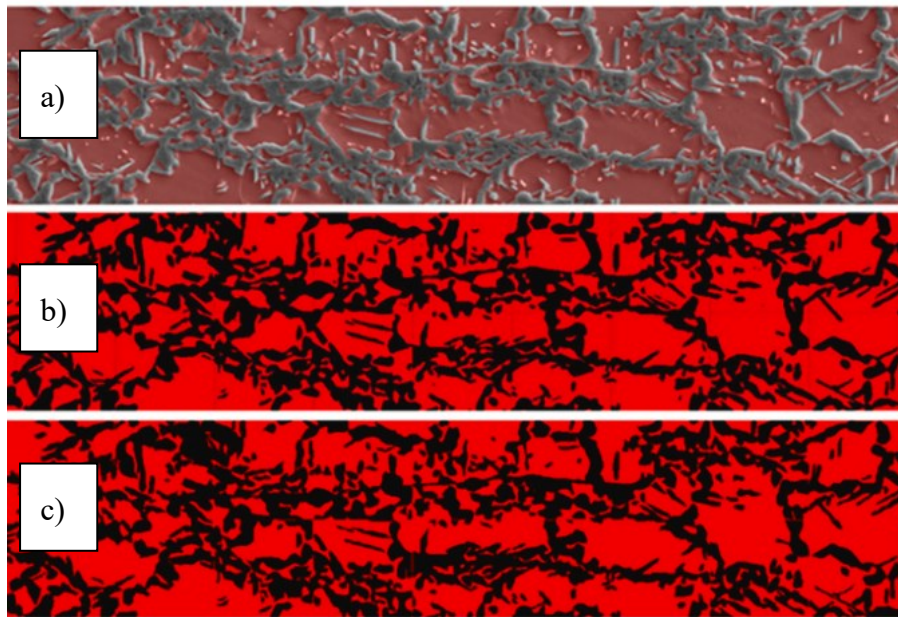
Fig. 2-5   a) Blend of SEM image and segmentation results, b) segmentation results, c) ground truth [15]

They have obtained an average of 93,1% PA and 88,2% IoU [15].

Luengo J et. al. recently published paper in which they tested different approaches and architectures. They were aiming to test different models of semantic segmentation for finding phases, precipitates, and defects of steel. They even tested unsupervised and semi-supervised deep learning methods, but the best performance had supervised learning. Unsupervised and semi-supervised learning approaches were not able to perform well, and they were not able to learn feature representations by themselves. They used nine different architectures for supervised deep learning: DeepLabv3, DeepLabv3+, FPN, Linknet, PAN, PSPNet, Unet, Unet++, PixelNet.

They used two different datasets to compare results. They used earlier mentioned UHCS [19] dataset and their own MetalDAM dataset, which was obtained with scanning electron microscope with 1280×895 and 1024×703 image resolution. These micrographs of steel have been generated from additive manufactured steel samples. They used data augmentation (vertical flip, horizontal flip, vertical and horizontal flip) for extending dataset size. Example of MetalDAM dataset is shown on Fig. 2-6 [2].
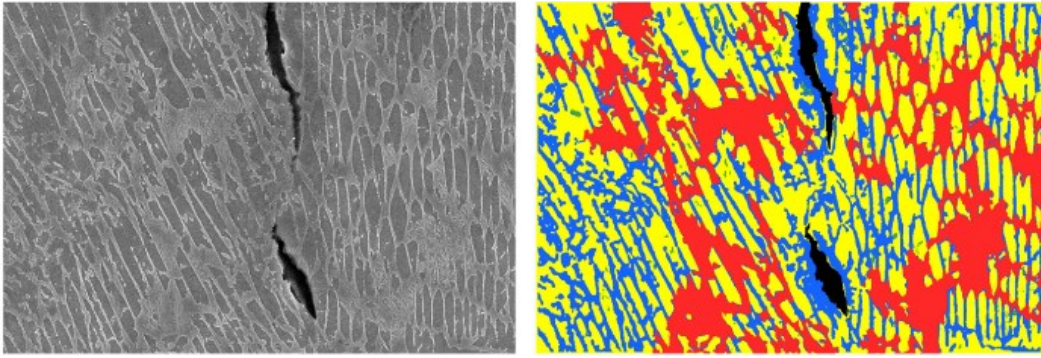
Fig. 2-6   Example of MetalDAM dataset. Left: original image. Right: labeled ground truth image [2].

They have concluded that U-net and U-net++ performed the best on tested datasets and they achieved average of 91,15% PA and 73,12% IoU with U-net++ architecture [2].

Nowroth C et. al., used deep learning semantic segmentation architecture DeepLabv3+ and software MATLAB for segmentation of microscopic images of laser beam welding seams for quantitative evaluation of welded contact. They have tested different model training parameters like type of optimization algorithm, amount of data used for training and image resolution [13].

They used optical microscopy to create their own large resolution micrographs for training. For labeling of 5 investigated classes, they used a tool Image Labeler which is a built-in tool in software MATLAB. After the training, they randomly split dataset into training, validation, and test set. Example of image pair is shown on the Fig. 2-7 [13].
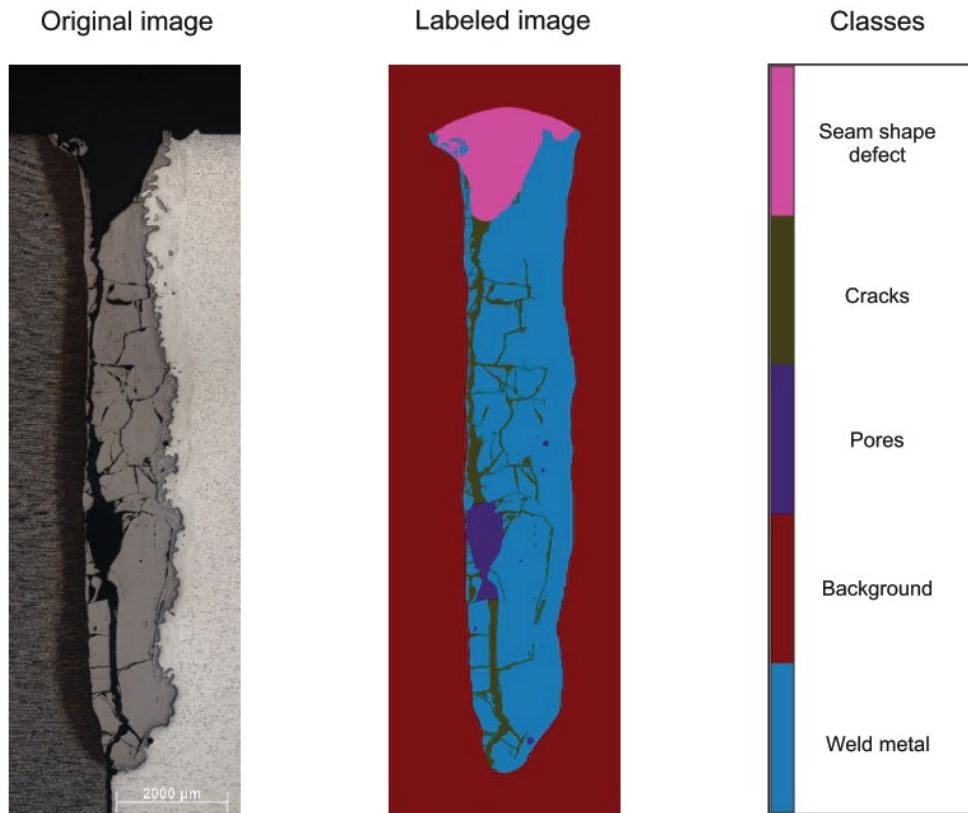
Fig. 2-7   Original image, labeled image and class differentiation [13].

They used only IoU for performance evaluation and they achieved average of 76,88% [13].

Performance evaluation of segmentation architectures

Evaluation of trained network is done on validating subset during the training of the network, and on the end, final model in evaluated on the test subset. First, segmentation map is created with trained network, and then it is compared with manually labelled ground truth image. For performance evaluation in Semantic Segmentation there are two main evaluation metrices commonly used. Generally, the higher value of chosen metric is, the network performs better.

▪ **Pixel accuracy (PA):** overall ratio of correctly predicted pixels.

$$PA = \frac{TP + TN}{TP + TN + FP + FN} \tag{1}$$

▪ **Region Intersection over Union (IoU):** Fraction of pixels that are correctly predicted to the union of pixels that are positive predications and belong to the target class. It punishes both overevaluation and undervaluation. For better understanding of IoU see Fig. 2-8.

$$IoU = \frac{TP}{TP + FP + FN} \tag{2}$$

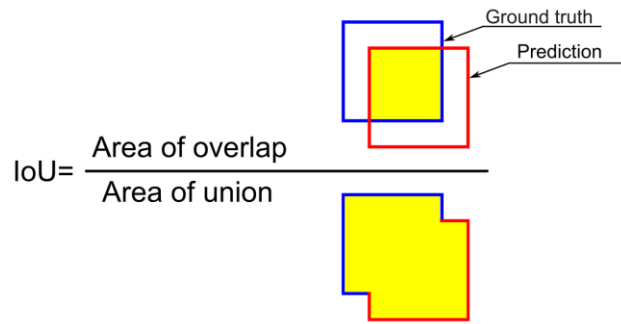$$IoU= \frac{\text{Area of overlap}}{\text{Area of union}}$$

Fig. 2-8   Region Intersection over Union

These metrices are calculated for each segmented class separately and then, overall metric is calculated as average of them. To simplify evaluation, mistake matrix is used, and her values are shown in Tab. 2-4.

Tab. 2-4   Values of mistake matrix.

Type	Meaning
TP - True positive	Number of correctly predicted pixels of segmented instance.
TN - True negative	Number of correctly predicted pixels of background.
FP - False positive	Number of incorrectly predicted pixels of segmented instance.
FN - False negative	Number of incorrectly predicted pixels of background.

Summarization of all semantic segmentation applications and their performances is shown in Tab. 2-5.

Tab. 2-5   Summarization of semantic segmentation applications and their performance.

Publication	Architecture	Dataset	PA [%]	IoU [%]
Shen C et. al. [14]	U-Net	DP	90.1 ± 5.8	83.2 ±9.3
		(Q&P)	85.4 ± 5.1	75.5 ± 6.5
De Cost et. al. [4]	Pixelnet - focal loss	UHCS	-	62.6 ± 2.5
	Pixelnet - crossentropy loss		-	75.4 ± 3.7
Roberts G et. al. [11]	DefectSegNet	Custom	96.61 ± 1.13	61.79 ± 2.13
Shen C et. al. [15]	U-Net		93.1 ± 3.1	88.2 ± 3.2
Nowroth C et. al. [13]	DeepLabv3+	Custom		76,88
Luengo J et. al. [2]	DeepLabv3	UHCS	87.76	62.15
		MetalDAM	81.03	56.69
	DeepLabv3+	UHCS	90.18	69.69
		MetalDAM	84.86	61.38
	FPN	UHCS	89.44	68.00
		MetalDAM	85.29	61.24
	Linknet	UHCS	88.7	67.75
		MetalDAM	84.69	55.36
	PAN	UHCS	88.4	64.68
		MetalDAM	84.44	58.24
	PSPNet	UHCS	88.2	63.33
		MetalDAM	79.09	53.44
	Unet	UHCS	91.15	72.45
		MetalDAM	86.33	59.95
	Unet++	UHCS	91.15	73.12
		MetalDAM	87.04	61.49
	Pixelnet	UHCS	90.77	70.79
		MetalDAM	84.07	51.41

2.2.3 Single image super-resolution

Super-resolution is used for many different tasks, but not often in microscopy and metallography yet. That's why I have reviewed few other papers that are focusing on development of GAN (Generative Adversarial Network) based and supervised approaches of image super-resolution [21].

Super-resolution architectures

Ledig C et. al created resolution enhancing architecture called SRGAN (Super-Resolution Generative Adversarial Network). Architecture was built from two different neural networks called: Generator and Discriminator. These two networks have been trained simultaneously against each other. Within training, they were comparing 3 different loss functions: MSE (mean square error), VGG22 and VGG54 loss [6].

They performed training on ImageNet dataset which contains 350 000 high resolution (HR) images. For validation of network performance, they used four different benchmark datasets already created for the purpose of development and researching image super-resolution architectures. They used Set5, Set14 and BSD100 dataset. All experiments were performed with 4× scale factor between low and high-resolution images. For evaluation, they used common image comparison metrices PSNR (peak signal-to-noise ratio) and SSIM (structure similarity index) which will be further explained in next section. Examples of results of the work are shown on Fig. 2-9. They managed to get 30,64 dB of PSNR and 0,8701 of SSIM with their best performing approach [6].

Fig. 2-9   Left to right, Bicubic Interpolation deep residual network, deep residual generative adversarial network, original high-resolution image. Corresponding PSNR and SSIM in brackets (4x upscaling) [6].

Wang X et. al. followed up on previous work and created architecture called ESRGAN (Enhanced Super-Resolution Generative Adversarial Network). They made modifications to this previous network by removing batch normalization layers and adding some other layers to the network [7].

28

They trained network on DIv2K, Flickr2K and Outdoor Scene Training datasets used for development models for image super-resolution. For evaluation of network performance, they used Set5, Set14, BSD100, Urban100 and PIRM datasets. For evaluation they used only PSNR. Example of results and comparison of their work with other state-of-the art methods is shown on Fig. 2-10. They were able to get highest PSNR from 20dB to 30dB depending on the images in dataset [7].
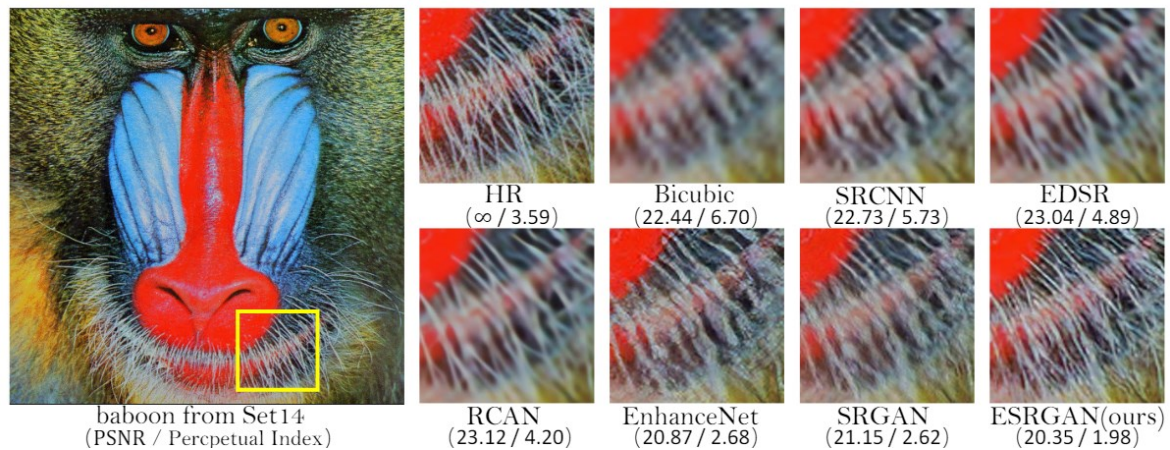

Fig. 2-10   Results of ESRGAN compared with different methods [7].

De Haan et. al. applied GAN based image super-resolution network with the aim of enhancing images made by scanning electron microscopy. They created their own architecture built from generator and discriminator [12].

They used scanning electron microscopy to create high resolution images. They used the 2× scale factor between low and high-resolution images. Training image pairs were created by taking high resolution images and then by lowering magnification of microscope and capturing low resolution (LR) image with same field of view example of result of their work is shown on Fig. 2-11 [12].
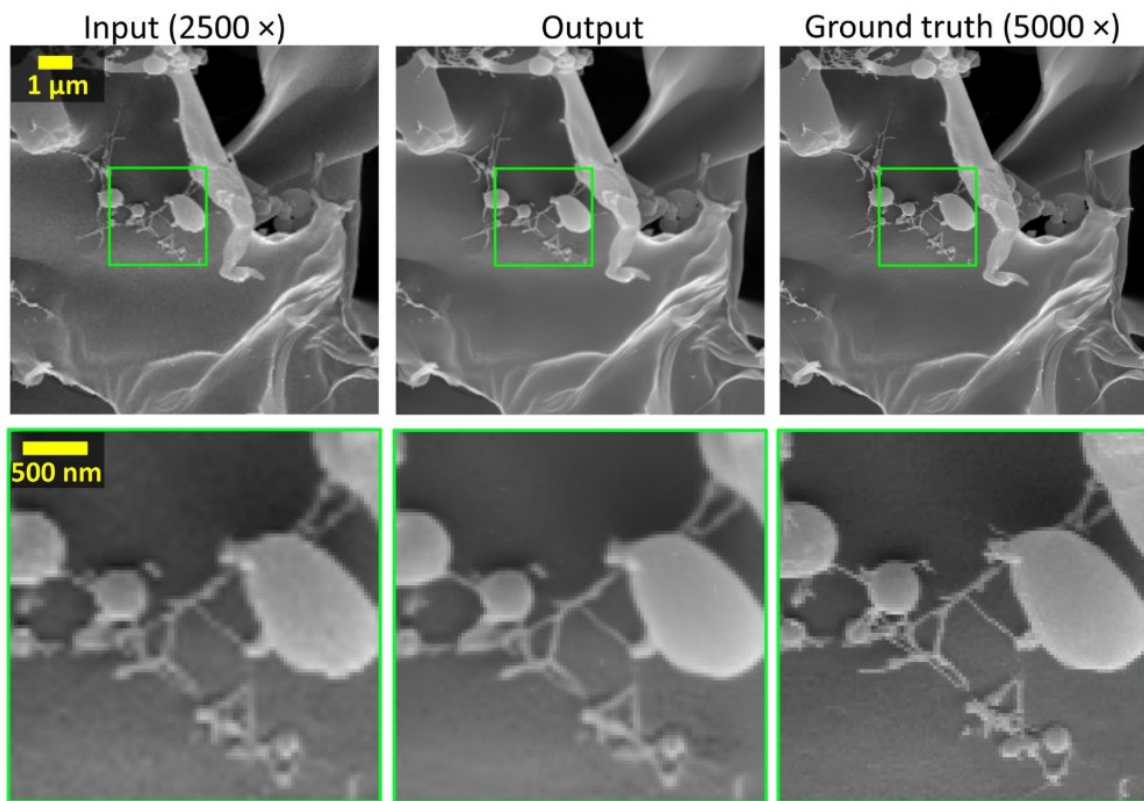
Fig. 2-11 Examples of up-scaled hydrogel. From left, input low resolution image, output of neural network, original high-resolution image [12].

Li et. al. used two types of image resolution enhancing architectures (SRGAN and SRResNet) to enhance images of biological cells. The images were obtained by laser confocal microscopy. These super-resolution techniques enhanced the boundaries of the image instances. They also clarify that they managed to enhance the mean fluorescence intensity which brought more embedded information in laser confocal images. Authors also clarify that with this method, high resolution SEM (scanning electron microscope) images can be taken faster with reduction both electron charging and damage to the samples [16].

They used 8294 images from confocal laser scanning microscopy. The dataset was split to a training set and test set in ratio 100/45. Result of their work is shown on Fig. 2-12. They have achieved 37.47 dB of PSNR and 0.94 of SSIM with SRGAN model [16].
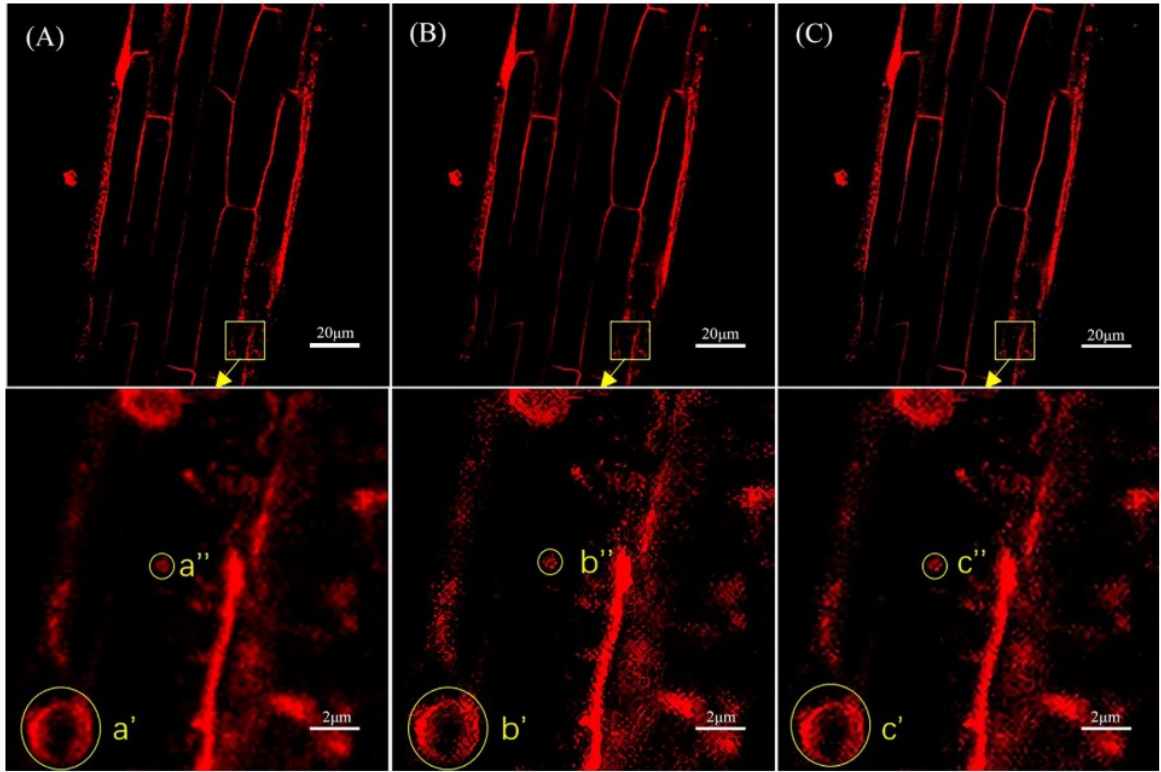
Fig. 2-12   (A) Original image (1024x1024 pixels) and local magnification (102x102 pixels), (B) SRGAN image (4096x4096 pixels) and local magnification (410x410 pixels), SRResNet image (4096x4096 pixels) and local magnification (410x410 pixels) [16].

Performance evaluation

For performance evaluation of image restoration, image super-resolution and in general for comparison of two images, there have been developed many different metrices: MOS PSNR, SSIM, IFC, LPIPS, NIQE, PIQE and NRQM [22]. In screened publications, PSNR, SSIM and MOS are mainly used.

- **PSNR (peak signal-to-noise ratio)**: This metric evaluates the pixel level intensity between original high-resolution image and generated high-resolution image. It is defined in decibels (dB) and higher value indicates less distortion and better quality of generated image. PSNR between reference image of ground truth f and test image g both of size WxH pixels is then calculated as [23]:

$$PSNR(f,g) = 10log_{10}(\frac{255^2}{MSE(f,g)}) \qquad (3)$$

where MSE is:

$$MSE(f,g) = \frac{1}{WH}\sum_{i=1}^{M}\sum_{j=1}^{N}(f_{ij} - g_{ij})^2 \qquad (4)$$

31

- **SSIM (structure similarity index)**: This metric does not evaluate pixel level but considers whole image region. It evaluates generated high-resolution image with original image in terms of structure, brightness, and contrast. It produces scores in range [–1, 1], where 1 represents a perfectly similar image. SSIM is defined as [23]:

$$SSIM(f,g) = l(f,g)c(f,g)s(f,g) \qquad (5)$$

where:

$$\begin{cases} l(f,g) = \frac{2\mu_f\mu_g + C_1}{\mu_f^2 + \mu_g^2 + C_1} \\ c(f,g) = \frac{2\sigma_f\sigma_g + C_2}{\sigma_f^2 + \sigma_g^2 + C_2} \\ s(f,g) = \frac{\sigma_{fg} + C_3}{\sigma_f\sigma_g + C_3} \end{cases} \qquad (6)$$

The first term in equation (5) is the luminance comparison function, which measures the closeness of the mean luminance of two images, denoted as μ_f and μ_g. The second term compares the contrast of two images where contrast is measured with standard deviation σ_f and σ_g. The third term is structure comparison function which measures correlation between two images. Constants C_1, C_2 and C_3 are positive numbers used to avoid division by zero [23].

- **MOS**: This metric is created as a score by observation of professionals in related fields who observe the pictures with their eyes. They give a score in range [1,5] where larger the number is, the better the quality of the image is [21].

Tab. 2-6 shows performance metrices of mentioned publications.

Tab. 2-6 Summarization of image super-resolution applications and their performance.

Publication	Model	Dataset	PSNR [dB]	SSIM	MOS
Ledig C et. al [6]	SRGAN	Set5	29.4	0.8472	3.58
		Set14	26.02	0.7397	3.72
		BSD100	25.16	0.6688	3.56
Wang X et. al. [7]	ESRGAN	Set14	26.4		
		BSD100	27.69		
Li et. al. [16]	SRGAN	Custom	37.47	0.94	3.89
	SRResNet		39.3	0.97	3.62

2.3 Review of selected neural network architectures

2.3.1 U-net for Semantic Segmentation

U-net architecture

Ronneberger et. al. first presented U-net architecture for segmentation of biological images. They managed to develop a network which can learn from only a small number of images. They used strong augmentation of initial small dataset and managed to outperform other segmentation architectures. Architecture is shown on Fig. 2-13 [18].

U-net architecture consists of two main parts. Contraction path (encoder) which captures context of features on image and expansion path (decoder) which enables localization of features. The encoder part consists of typical convolutional layers each followed by reflected linear unit (ReLu) and max pooling operation. Each down-sampling step, the number of feature channels is doubled. In the decoder side of U-net, each step consists of up-sampling of features followed by convolution so called up-convolution, which reduces the number of feature channels by half. Each step of the decoder is connected to the encoder using a cropped feature map from the encoder. On the final output layer, convolution is used to map each feature to desired number of segmented classes [18].
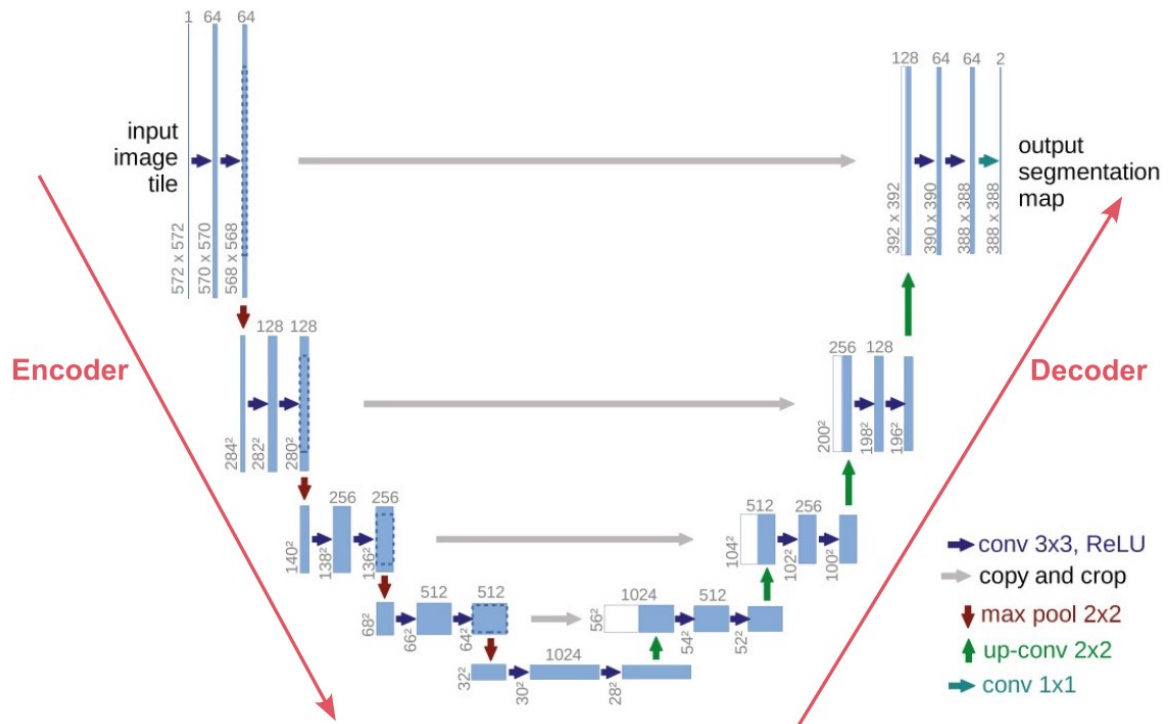

Fig. 2-13 The U-net architecture, features blue boxes that represent multi-channel feature maps. The number of channels is noted above each box, and the x-y dimensions are displayed at the lower left corner. White boxes indicate copied feature maps, and arrows show the different operations [18].

Training loss functions

As researchers pointed out in section 2.2.2, it is important to try different loss functions during the training. Also, Jadon S recently reviewed loss functions for semantic segmentation tasks. Loss functions in general provide a measure of how well the model's predictions match the expected outputs, guiding the training process through optimization. I have chosen two widely used loss functions for training of segmentation architectures also mentioned in 2.2.2. I have chosen categorical cross entropy loss function and categorical focal loss function [24].

- **Categorical cross entropy loss function:** It aims to minimize the difference between the predicted probability and the actual label for each pixel across all classes. It is ideal for tasks where each pixel belongs to one of several classes, and the classes are mutually exclusive.
- **Categorical focal loss function:** It focuses more on hard-to-classify pixels (where the model is unsure) and less on easy cases where the model is already performing well. It is useful in scenarios where there is a significant class imbalance, such as in medical imaging or metallographic imaging [25].

2.3.2 SRGAN for image Super-Resolution

From section 2.2.3 we have chosen SRGAN architecture for experimenting in this thesis. This architecture is based on GAN networks which consist of generator network and discriminator network. These two networks are trained together during the training process. The generator network is trying to generate new images from input low resolution image, and discriminator then tried to recognize if the generated image is real or fake. This means that generator is trying to learn how to fool discriminator and discriminator is trying to learn how to recognize if the generated image is real or fake. This approach is shown on schematics on Fig. 2-14. After every training iteration, trainable parameters of networks (weights and biases) are updated. After training, only generator network is used for high resolution image generation [21].
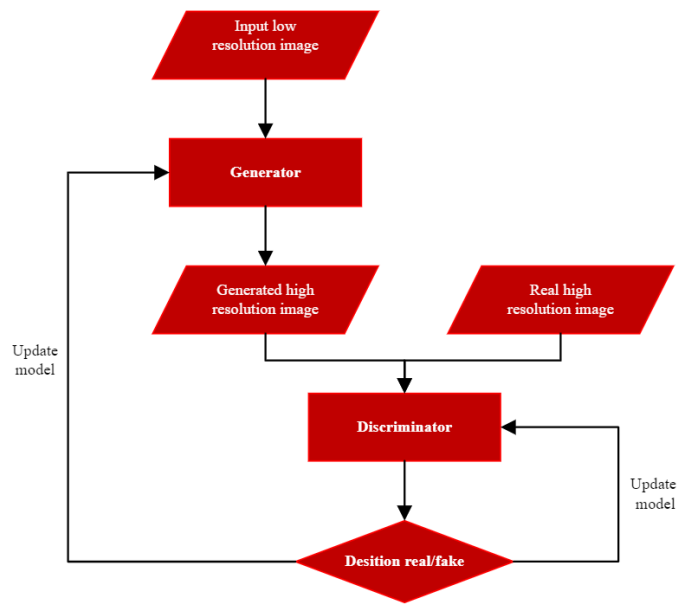
Fig. 2-14   Schematics of training process of GAN based image super-resolution neural networks.

Structure of SRGAN architecture proposed by Ledig et. al is shown on Fig. 2-15.
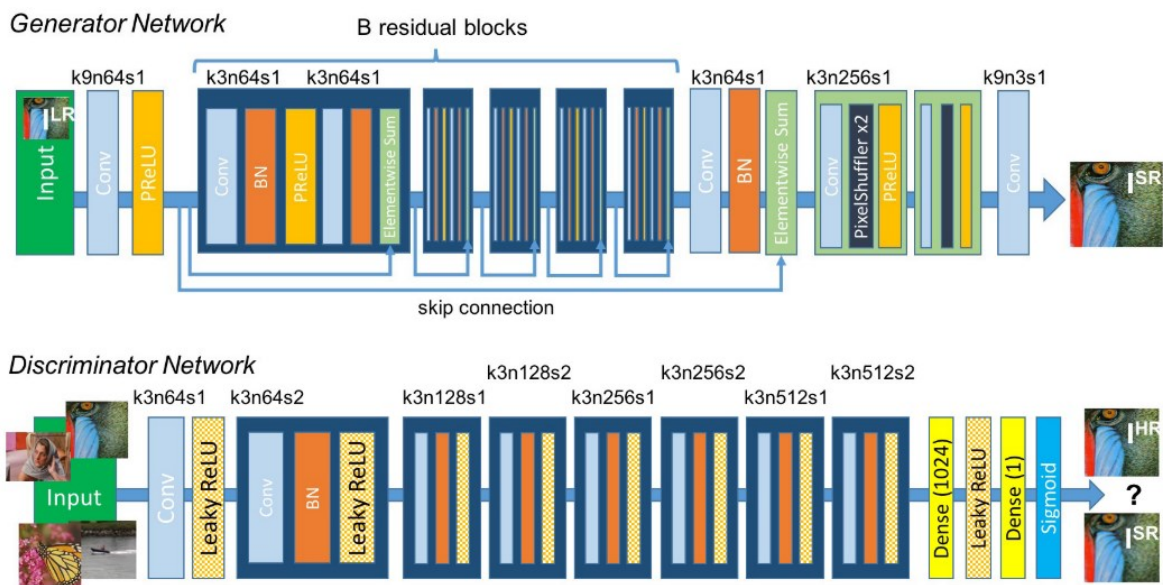

Fig. 2-15   Generator and Discriminator Network of SRGAN with corresponding kernel size (k), number of feature maps(n) and stride (s) indicated for each convolutional layer [6].

2.4   Systematic review summarization and knowledge gap

In this section, we will focus on addressing the research questions posed in Section 2.1.

Q1 - What methods are used for evaluating metallographic cross sections made by additive manufacturing?

As has been said, Murr L reviewed recent progress in metallography of additive manufactured parts. It follows from methods of referring articles that no one used methods based on neural networks for evaluating additively manufactured material for porosity types [10].

At the same time, DebRoy T et. al. pointed out, that machine learning and neural networks generally could be used at all steps of metal printing, not only for evaluating cross-sections. It has strong potential as a tool to predict microstructures, properties, and defects during 3D printing process [9].

Q2 - Are there any applications of Semantic Segmentation or Super-Resolution in metallography?

Recently, there have been a lot of different applications for Semantic Segmentation and image Super-Resolution with deep learning in different industries. Researchers have developed different models that are suitable for various tasks and for diverse amounts of training data.

It was found out that there are some applications of neural networks in metallography, not only in image processing of metallographic cross sections, but for example to computationally design materials that meet target property requirements. To evaluate cross-sections, researchers were trying to use multiple classification methods. Semantic Segmentation is suitable for pixel level accuracy, and it has been used in many approaches to evaluate metallographic cross-sections, even for materials made by additive manufacturing.

All publications that presented Semantic Segmentation for categorizing structures of steels testified, that if they used the right architecture with right loss function, then training process required only small experimental dataset.

In the domain of image Super-Resolution of metallographic images, there are no publications of applications in additive manufacturing yet. Two publications were found which discuss application of Super-Resolution in electron microscopy and confocal laser scanning microscopy. De Haan et. al. pointed out, that image Super-Resolution is a powerful and practical tool to computationally improve resolution of microscopic images [12].

Generally, for image processing tasks, researchers claim that methods based on neural networks are smarter and more flexible than traditional image processing methods, but at the same time, they demand large computing power.

Q3 - What types of neural network architectures are used for image Semantic Segmentation?

As it has been said, there are three main approaches to deep learning in semantic segmentation: deep supervised learning, deep unsupervised learning, and deep semi supervised learning. We mainly aimed for publications that were focused on semantic segmentation with supervised learning in metallography. As it has been shown in Tab. 2-5, researchers have tested different architectures with multiple loss functions to this task.

Luengo J et. al. recently tested and compared different architectures on two datasets. They pointed out that Unet++ and Unet architectures achieved the best results in both UHCS and MetalDAM dataset, so they stated, that these models are robust and effective in metallographic domain. At the same time, it is important to train networks with different loss functions to test which works the best way on trained dataset.

Q4 - What types of neural network architectures are used for image Super-Resolution?

In past years, image super-resolution has been used in many different applications. It has been used mainly to enhance video resolution and general single image resolution. There are many different approaches to this task [5]. Researchers have shown that GAN based networks perform best for recovering edges of instances. This is a strong property of GAN based Super-Resolution architectures, that is why we focused on reviewing recently developed architectures.

2.5 Interpretation of critical review results for this thesis

From critical review we can see that usage of neural networks for evaluation of microscopic images has a strong potential. In evaluation of cross sections in 3D printed material domain, there has not been many applications and experiments performed yet.

Such a use case of evaluation of cross sections to obtain porosity levels and types of porosities, nobody studied before. Usage of neural networks in evaluation process could bring more information from the material and further enhance process parameters of 3D printing.

In the domain of image segmentation, applications in microscopic imaging have already been studied before. I must point out that researchers have shown that every use case needs its own training and optimization process of neural network model. To this thesis, I also need to train the model by myself and find optimal training parameters for it. From many available neural network architectures, I have chosen U-net for simplicity and because there is evidence that this model performs well on metallographic images [2, 14, 15].

About image Super-Resolution, I must point out that it is a really new approach in domain of microscopic images. It has not been much studied in combination with segmentation techniques. There is strong potential in research in this field due to the achievement of subpixel accuracy with further image processing. Thus, for the purpose of my thesis, I have focused on usage of SRGAN architecture, and explored capabilities of this model in the field of metallography [16].

3   PROBLEM ANALYSIS AND GOALS OF THESIS

3.1   Problem definition

For evaluation and optimization of process parameters during research and development of new metallic materials made by powder bed fusion different parameters are used (laser power, scanning speed, scanning strategy etc.). These are nowadays still estimated from megalographic cross sections. There is one single parameter "porosity" evaluated from cross sections and simple image processing methods like threshold-based segmentation is used for this evaluation. Threshold-based segmentation is often sensitive to noise, imperfections, or more complicated textures of cross-sections. This can lead to inaccurate results of segmentation. On Fig. 3-1, the most typical imperfections which are often problem during segmentation can be seen. On the example there is dried cleaning medium (ethanol or isopropyl alcohol) which can dry around larger pores during cleaning of specimens, entrapped powder inside of pores which is not holding with base material, or light of microscope illumination which reflects from inside of pores. All these imperfections bring inaccuracy to the results of segmentation.
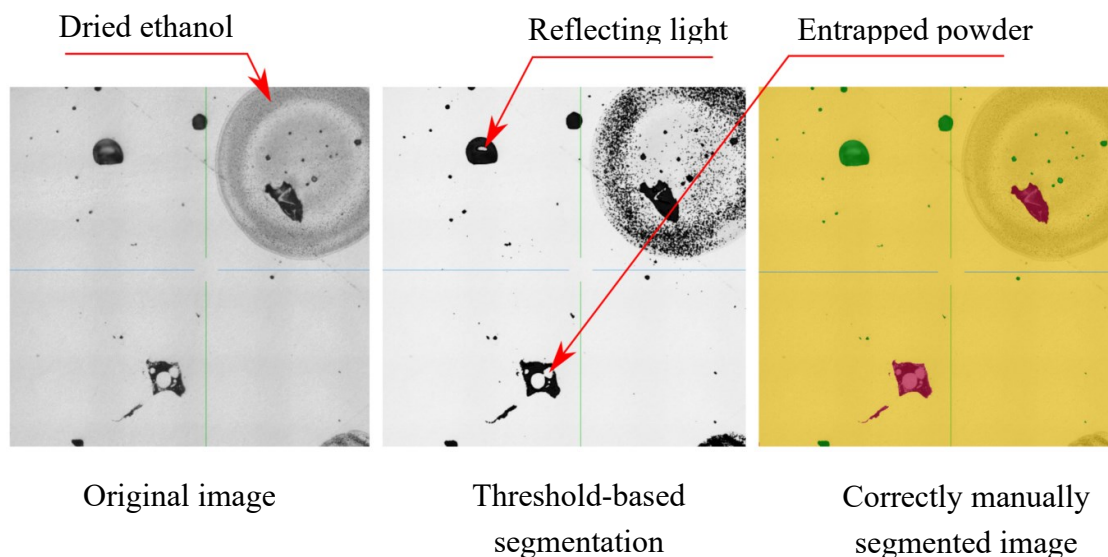

Fig. 3-1   Example of imperfections which can occur during image segmentation.

There are two main types of defects which can occur inside of material.

- **Lack of fusion porosity (LoF)** – This type of porosity occurs inside of material due to the combination of low laser power and fast scanning speed. Structures of badly fused material or even non fused powder can be seen on metallographic cross section.

- **Gas porosity (Gas)** – This type of defect similarly occurs in material, but the reason is combination of high laser power and slow scanning speed. Gas porosity has a spherical shape. Gas porosity can also occur due to humidity evaporation during printing process.

Parameter "porosity" does not distinguish between those types of defects so this parameter could be often insufficient during evaluation. There would be suitable to develop a method which would be able to automatically distinguish between those types of defects to get more precise results and to deeper understand what is happening inside of the material.

Simultaneously, there exists potential in employing an image upscaling neural network algorithm to artificially upscale and enhance the edges and structures of features under study in images. This process would provide additional data for segmentation. The idea is that an increase in data should bring more robust and accurate segmentation results.

3.2 Goals of thesis

The main goal of the thesis is to use advanced image processing methods based on deep neural networks (image semantic segmentation and super-resolution) for the processing and analysis of microscopic images of metallographic cross-sections of materials made by powder bed fusion.

The specific objectives are:

- Selection of suitable "off-the-shelf", ready-made neural network architectures for super-resolution and semantic segmentation tasks.
- Creation of a custom dataset which includes metallographic images of three 3D printed materials (AlSi10Mg, CU99, and Ti6Al4V) with manually labelled defects, suitable for training super-resolution and semantic segmentation neural networks.
- Development of custom training loop for training of chosen neural network models and fine-tuning of training parameters with the goal of achieving the best-performing neural network models for specific image super-resolution and semantic segmentation task on created dataset.
- Determination of whether there is a positive influence on the results of semantic segmentation by enhancing its input image with a super-resolution neural network.
- Investigation of whether a semantic segmentation network trained exclusively on a single material (AlSi10Mg) can still perform robust image segmentation of required defects on images of different materials (CU99 and Ti6Al4V) that it has not previously encountered.

- Implementation of the most effective approach in custom solutions for defect analysis of cross-sections in powder bed fusion research of new materials, to speed up defect analysis and categorize defects of metallographic cross-sections.

3.3 Research questions

- **Question 1.**: What will happen if we train segmentation network on AlSi10Mg and let it perform image segmentation on different material?
- **Question 2.**: What will happen with results of image segmentation if we enhance its input image with image up-scaling algorithm?

3.4 Hypothesis

Hypothesis 1.: If image segmentation neural network model will be trained only on material AlSi10Mg, it will still be able to perform accurate segmentation of defects on material Cu99 and Ti6Al4V because of defects shape is simar cross all material types.

Hypothesis 2.: If an image up-scaling algorithm based on GAN is used before image segmentation of metallographic cross sections, then the results of measured performance metrices will have higher value than the results of stand-alone segmentation. This hypothesis has been obtained due to the latest research in usage of image super-resolution in combination with segmentation tasks by Frizza et. al. [26].

4 MATERIAL AND METHODS

4.1 Methodology

The first step was to create images of given metallographic cross sections for each material (AlSi10Mg, Cu99 and Ti6Al4V). Next step was to manually label defects (lack of fusion porosity, gas porosity and base material) on images and create image pairs (original image and its labels) which were used for image segmentation model training. For image up-scaling model only original images without labels were used. After dataset preparation, neural network architectures were trained. On Fig. 4-1 and Fig. 4-2, there is shown methodology of this work.

To answer the first question, segmentation model was trained only on images of AlSi10Mg. Then the model was tested on images of Cu99 and Ti6Al4V.

To answer the second question, upscaling-model and segmentation model were trained on images of AlSi10Mg. Then upscaling model and segmentation model were serially connected and compared with standalone segmentation model.

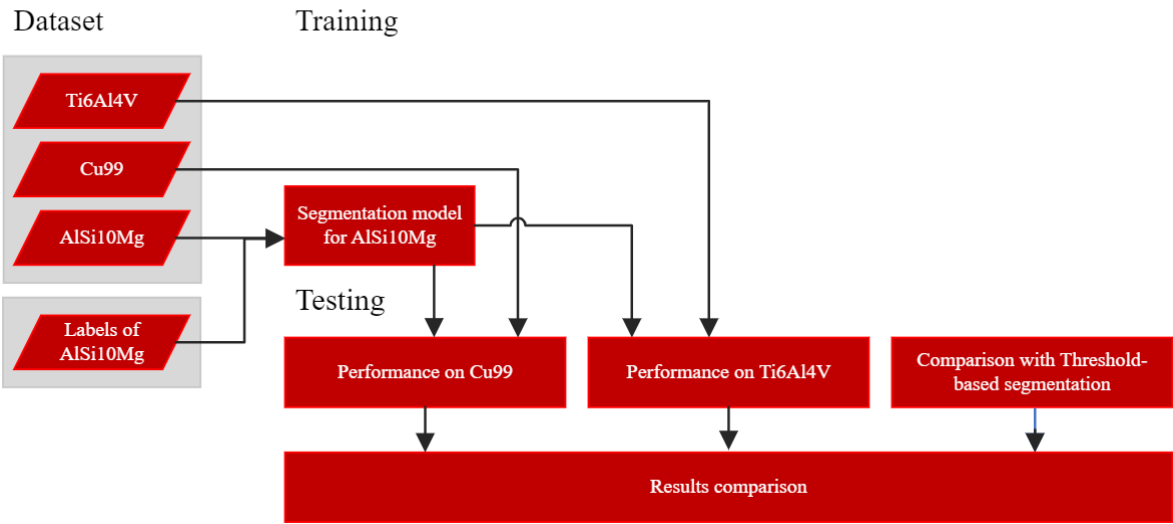The best performing approach was then implemented in custom software solution.
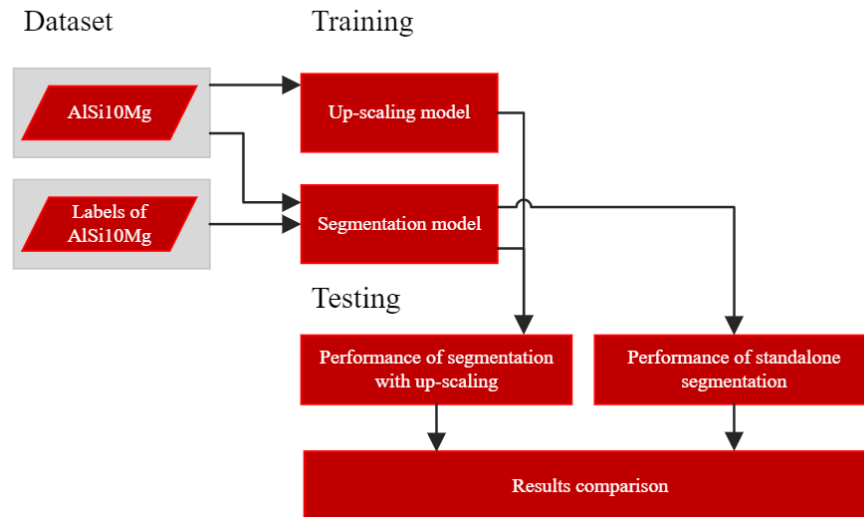

Fig. 4-1   Methodology question 1

Question 2

Fig. 4-2   Methodology question 2

4.2 Instrumental equipment

4.2.1 Microscope Keyence VHX-6000

Digital Microscope Keyence VHX-6000 is equipped with lens Z250R [27] (Fig. 4-3). This lens is able of 250x–2500x magnification. For this work and dataset production, magnification of 1000x was used. This magnification was selected due to limited microscope setup conditions, so the microscope images were not blurred and features on the image were still recognizable due to outside physical factors (vibrations of the table during imaging). Images of metallographic cross-sections were created by stitching function, so multiple pictures were stitched together. Final exported images had size of 9332 x 9317 pixels, and they were exported in lossless .tiff format.

Fig. 4-3   Digital microscope Keyence VHX-6000[27].

4.2.2 Computer setup

In Tab. 4-1 there is PC hardware setup which was used during software development, dataset labelling and neural networks training.

Tab. 4-1   PC setup

Parameter	Value
Processor	13th Gen Intel(R) Core (TM) i5-13600KF (20CPUs) 3.5GHz
Graphics processing unit	NVIDIA GeForce RTX 3060 Ti (8 GB memory)
Operational memory	64GB

Graphics processing unit (GPU) NVIDIA GeForce RTX 3060 Ti with 8GB of memory is the most crucial component for neural network training, and it has computing power of 8,6 according to computing capability rating by Nvidia [28].

4.3 Dataset

4.3.1 Initial image preprocessing.

Custom datasets consist of cross-sections of AlSi10Mg, Cu99 and Ti6Al4V made by powder bed fusion. There were 6 images created for each material with mentioned microscope with 1000x magnification. Each image displayed different process parameters which were used during 3D printing to have variety of defect types shown on the images. Examples of microscope output images are shown on Fig. 4-4. Each image was then converted from 3 channel RGB (red, green, blue) to only single black and white channel (gray scale) and divided into smaller images with size of 4616 x 4608 pixels for further dataset labeling.
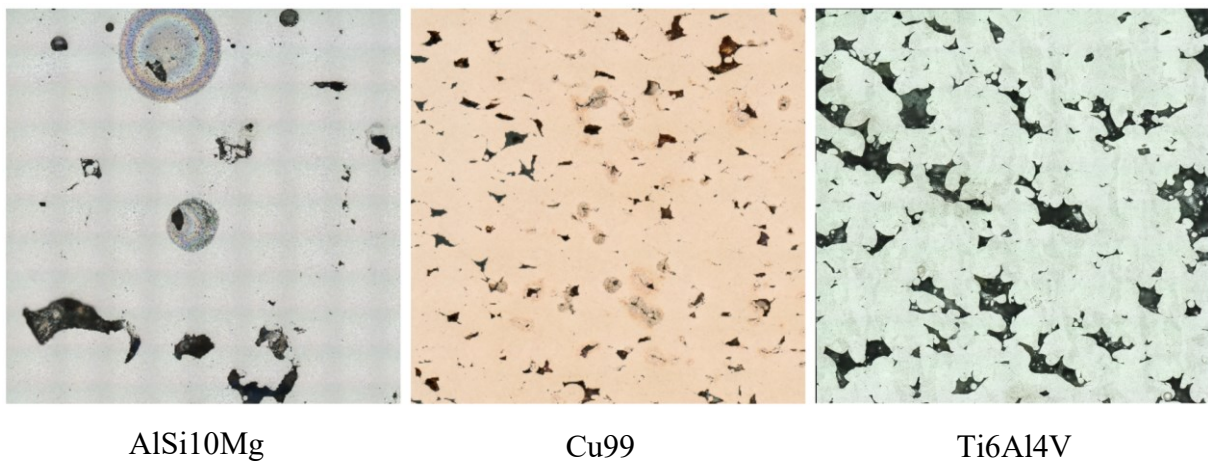

| AlSi10Mg | Cu99 | Ti6Al4V |

Fig. 4-4   Example of microscope output images (9332 x 9317 pixels)

4.3.2 Dataset for U-net training

For training of semantic segmentation model U-net, only images of AlSi10Mg were used. I have picked 14 images of AlSi10Mg from dataset (each 4616 x 4608 pixels), each representing different defect shapes and process parameters for diversity. Each smaller image was then manually annotated on pixel level with MITK [20] which is tool able of semiautomatic image labelling and uses various classic image processing algorithms to speed up labelling process. Region growing algorithm with combination of threshold-based segmentation method was used during annotation of defects. Three labels were created: lack of fusion porosity, gas porosity and base material. Examples of image labelled pairs are shown on Fig. 4-5.
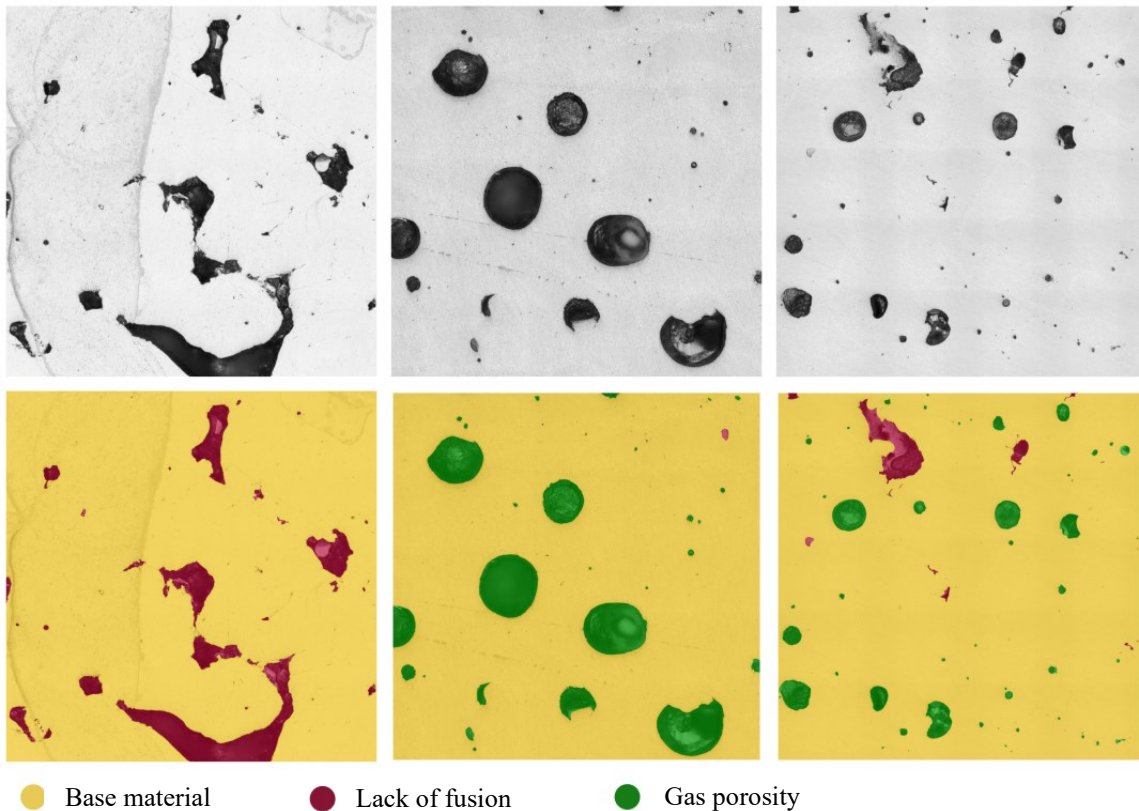
Fig. 4-5   Examples of labelled image pairs of AlSi10Mg (4616 x 4608 pixels).

On Tab. 4-2 there is ratio of each class annotated class in AlSi10Mg dataset. Dataset was imbalanced due to high number of pixels corresponding to base material.

Tab. 4-2   Class distribution through AlSi10Mg in dataset

Class	Ratio (%)
Base material	92,56
Lack of fusion porosity	4,26
Gas porosity	3,18

Dataset division

To correctly calculate performance metrices of the trained U-net, I had to create separate subsets of images to train, validate and test the network on the data which the network never seen during the training.

- **Training set** – Used to train the models and adjust the parameters of trained network.
- **Validation set** – Used for checking model performance and calculation of performance metrices during the training.
- **Testing set** – Used for final model performance evaluation. And on testing on materials which model had not seen before.

Then 11 images were taken from AlSi10Mg dataset and split to two subsets for training and validation in ratio 85/15. The remaining 3 images from dataset were picked for testing and I have ensured that these three images were of material process parameters which were not included in training and validation subset. For further performance testing and validation of hypothesis 1. on Cu99 and Ti6Al4V, I have similarly labeled 2 images (each 4616 x 4608 pixels) for each material (Cu99 and Ti6Al4V).

Dataset augmentation

Before neural networks training, the training set for U-net was artificially augmented. This is commonly done during the training of neural networks to prevent overfitting and to scale the size of dataset. Overfitting means that the network could potentially learn too well on the training data and perform poorly on the validation and training data.

I have augmented the dataset by commonly used augmentation techniques mentioned in 2.2.2. I have used random rotation by 0.5π, π, 1.5π, random horizontal and vertical flip, random brightness shift in range $\pm10\%$ and scaling in range [1, 1.5].

Before feeding the training data to the U-net training process, training images with their corresponding labels were downscaled by the factor of 4 because of graphics card low memory issues during the training. So for training and evaluation, downscaled images were used. Then I created a data augmentation generator module, which was then feeding the augmented data to the training process. This generator module took one image and its corresponding label, then randomly applied combination of mentioned augmentation techniques to the image pairs and cropped image on random location to corresponding U-net input image size (HxW) which was further passed to the training process. With this augmentation data generator, I have ensured that during the training, every training epoch (training iteration) there will be slightly different images shown to the network.

4.3.3 Dataset for SRGAN training

For experiments with SRGAN image super resolution neural network, I have used only high-resolution images of AlSi10Mg. Dataset for SRGAN training was randomly split in to two subsets similarly as with semantic segmentation training to train set and validation set in ratio 80/20.

For dataset augmentation I have used random cropping, random rotation by 0.5π, π, 1.5π, random horizontal and vertical flip, random brightness shift in range $\pm10\%$.

4.4 Neural networks models

For implementation of selected neural network models, python programing language was used with TensorFlow [29] platform, which is one of commonly used frameworks for implementation of neural networks and for work with matrix and tensor computation within GPU. All used libraries were installed to Anaconda environment for easier libraries version management [30]. This setup was running on the mentioned hardware with Linux Ubuntu operating system.

4.4.1 Segmentation model U-Net

U-Net architecture was picked from many suitable image segmentation architectures as it has been mentioned in section 2.5. For my use case it was modified to expect single channel gray scale image on its input and return three channel prediction maps for each segmented class (Lack of fusion porosity, Gas porosity, Base material). Also, I have been experimenting with the depth (number of trainable parameters) of the U-net. Since I was working with only single channel image, there were 3-times less input parameters for the network compared to RGB image, so each module of U-net did not have to have that many parameters. This was also done by other researcher's mentioned in section 2.2.2. Depth of U-net was one of the variable experimental parameters which I represented in depth multiplication factor δ. The structure of edited U-net used in this work is shown in Fig. 4-6.
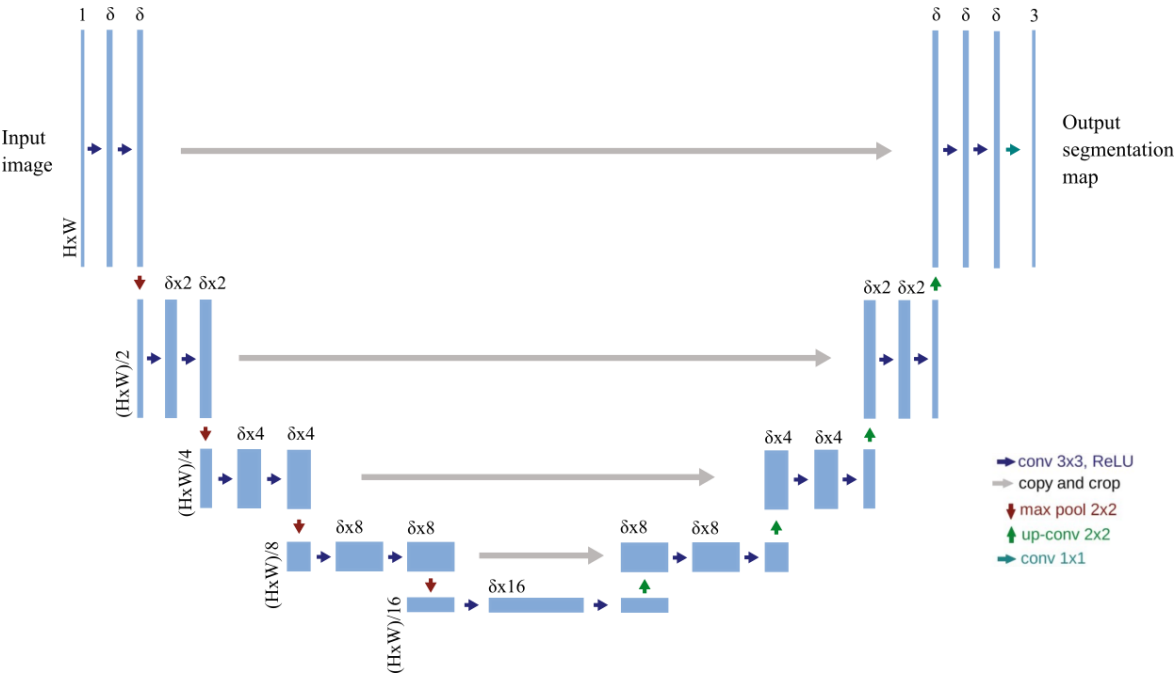

Fig. 4-6   U-net structure with custom input/output image size and custom depth multiplication factor δ.

4.4.2 Image super-resolution model SRGAN

As has been mentioned before, ready-made, off-shelf architecture SRGAN was used for experiments in this thesis.

4.5 Performance evaluation

4.5.1 Segmentation model evaluation

For evaluation of semantic segmentation model during the training, PA and IoU metrices were used which were reviewed in section 2.2.2. During experiments, I focused on PA and IoU metrices.

After fine tuning models and finding the best possible and suitable approach to segmentation, the model was also compared with threshold-based segmentation. For finding the threshold value, Otsu's method was used [31].

4.5.2 Image super-resolution model evaluation

I have evaluated the model during training with two of the mentioned performance metrices (2.2.3). PSNR and SSIM were used. As has been also shown from the systematic review, these two metrices are suitable for evaluation of two images to score structure similarities between two images.

4.6 Implementation.

The best performing model has been implemented in custom software solution. From classification maps which were output from segmentation, there have been computed percentual ratio of each type of defects compared to the base material.

5 RESULTS

5.1 Semantic segmentation

5.1.1 Selection of variable parameters for experiments

After initial experimenting with U-net training I have realized that different setup of training process brings different PA and IoU values of resulting neural network. Because of this I have designed an experiment for finding the most suitable parameters for the given dataset and dataset augmentation setup.

I have selected four variable parameters, which as observed in both my initial findings and reviewed papers, influence the resulting metrics of IoU and PA for neural network models on unseen data. I was experimenting with two types of loss functions (categorical cross entropy loss and categorical focal loss), two types of neural network input and output image patch sizes, depth of neural network model (number of trainable parameters) represented in depth multiplication factor δ and lastly, batch size (number of training examples processed in one forward and backwards pass through the network during the training process). Selected variable parameters and their values are shown in Tab. 5-1.

Tab. 5-1   Selected variable parameters for experiments.

Loss function	Input/ output patch size (H x W) [pixels]	Depth multiplication factor δ [-]	Batch size [-]
Cross entropy loss	512x512	2 (30 537 trainable parameters)	4
Focal loss	1024x1024	4 (121 663 trainable parameters)	8
		8 (485 691 trainable parameters)	16
		16 (1 940 851 trainable parameters)	32

5.1.2 Training experiments

To train a U-net architecture, a custom iterative training experiment was developed. The training loop was designed to iterate through all variable parameters mentioned in Tab. 5-1 and save the model for each parameter combination to the specified folder. The commonly used Adam optimizer [32] was used to optimize the network parameters. During the training process, the validation loss was plotted, and the model was saved each time the validation loss decreased. Only the models with the lowest loss were retained from the training process.

A test set of images was used to evaluate each trained model. A testing loop was developed, which iteratively loaded each saved model and assessed its performance. The performance metrics of IoU and PA were then written in a .json file for further evaluation. PA was calculated for each class, and the values of PA and mean values of IoU are presented on heat maps on Fig. 5-1 and Fig. 5-2. Experiments which exceeded the limit of GPU memory are marked as OOM (out of memory).

Categorical cross-entropy loss								
Batch size	4		8		16		32	
(HxW)	512x512	1024x1024	512x512	1024x1024	512x512	1024x1024	512x512	1024x1024
Depth multiplication factor δ — 2	77,24%	69,74%	66,29%	68,25%	71,00%	70,88%	37,44%	OOM
4	80,66%	70,82%	72,42%	78,97%	79,95%	OOM	82,59%	OOM
8	77,46%	75,86%	77,16%	OOM	80,06%	OOM	80,64%	OOM
16	80,01%	67,84%	77,20%	OOM	OOM	OOM	OOM	OOM
Categorical focall loss								
Batch size	4		8		16		32	
(HxW)	512x512	1024x1024	512x512	1024x1024	512x512	1024x1024	512x512	1024x1024
Depth multiplication factor δ — 2	71,77%	69,06%	75,29%	63,29%	41,07%	65,59%	69,36%	OOM
4	70,68%	65,65%	75,81%	66,87%	81,39%	OOM	72,77%	OOM
8	72,43%	75,15%	76,17%	OOM	73,11%	OOM	76,03%	OOM
16	80,43%	70,48%	76,50%	OOM	OOM	OOM	OOM	OOM

Fig. 5-1　　Mean IoU values of trained models from experiment.

Categorical cross-entropy loss

Batch size			4						8						16						32					
(HxW)			512x512			1024x1024			512x512			1024x1024			512x512			1024x1024			512x512			1024x1024		
Class			LoF	Gas	Base	LoF	Gas	Base	LoF	Gas	Base	LoF	Gas	Base	LoF	Gas	Base	LoF	Gas	Base	LoF	Gas	Base	LoF	Gas	Base
Depth multiplication factor δ	2		88,4%	69,3%	99,7%	78,8%	78,4%	99,5%	73,4%	62,0%	99,7%	78,4%	57,9%	99,6%	90,0%	48,1%	99,5%	87,0%	61,6%	99,3%	3,3%	61,8%	99,8%	OOM	OOM	OOM
	4		90,8%	79,7%	99,6%	76,7%	78,6%	99,7%	83,5%	72,4%	99,6%	87,0%	78,6%	99,6%	91,6%	68,8%	99,5%	OOM	OOM	OOM	92,8%	77,8%	99,4%	OOM	OOM	OOM
	8		86,4%	73,2%	99,8%	88,6%	77,9%	99,5%	86,9%	71,0%	99,7%	OOM	OOM	OOM	89,5%	72,7%	99,7%	OOM	OOM	OOM	91,1%	67,4%	99,5%	OOM	OOM	OOM
	16		92,6%	66,2%	99,6%	78,2%	68,8%	99,5%	87,9%	73,6%	99,7%	OOM	OOM	OOM	OOM	OOM	OOM	OOM	OOM	OOM	OOM	OOM	OOM	OOM	OOM	OOM

Categorical focal loss

batch size			4						8						16						32					
(HxW)			512x512			1024x1024			512x512			1024x1024			512x512			1024x1024			512x512			1024x1024		
Class			LoF	Gas	Base	LoF	Gas	Base	LoF	Gas	Base	LoF	Gas	Base	LoF	Gas	Base	LoF	Gas	Base	LoF	Gas	Base	LoF	Gas	Base
Depth multiplication factor δ	2		80,8%	56,9%	99,8%	79,9%	62,8%	99,5%	85,3%	64,7%	99,8%	82,7%	36,6%	99,3%	18,9%	35,9%	99,8%	72,5%	64,1%	99,5%	88,0%	31,3%	99,9%	OOM	OOM	OOM
	4		81,6%	75,0%	99,4%	70,6%	68,5%	99,7%	86,1%	58,8%	99,8%	70,9%	77,7%	99,7%	90,8%	64,9%	99,6%	OOM	OOM	OOM	81,7%	72,9%	99,7%	OOM	OOM	OOM
	8		80,6%	69,6%	99,8%	91,0%	78,7%	99,1%	87,1%	64,0%	99,7%	OOM	OOM	OOM	83,2%	67,2%	99,8%	OOM	OOM	OOM	85,4%	76,4%	99,7%	OOM	OOM	OOM
	16		89,0%	74,0%	99,8%	80,8%	65,0%	99,6%	86,9%	74,2%	99,6%	OOM	OOM	OOM	OOM	OOM	OOM	OOM	OOM	OOM	OOM	OOM	OOM	OOM	OOM	OOM

Fig. 5-2 Heat map for per class PA of trained models from experiment.

From the first look on the results, it is observed that categorical cross entropy loss leads to higher values of PA and IoU cross all tested training parameters. That is why I have focused on models trained on categorical cross entropy loss for further evaluation. With increasing batch size in combination with model depth, the limit of GPU memory is exceeded leading to OOM errors. Also, it is clear from Fig. 5-2 that models have hard time to classify gas porosity class.

On Fig. 5-3 there is Graph which shows relation of IoU to the model depth multiplication factor δ for different input/output image patch sizes. This was for all performed experiments. We can see that IoU metric is lower for higher input/output image patch size. The best IoU values are for model with model depth multiplication factor δ of 4 and input/output image size of 512x512 pixels.
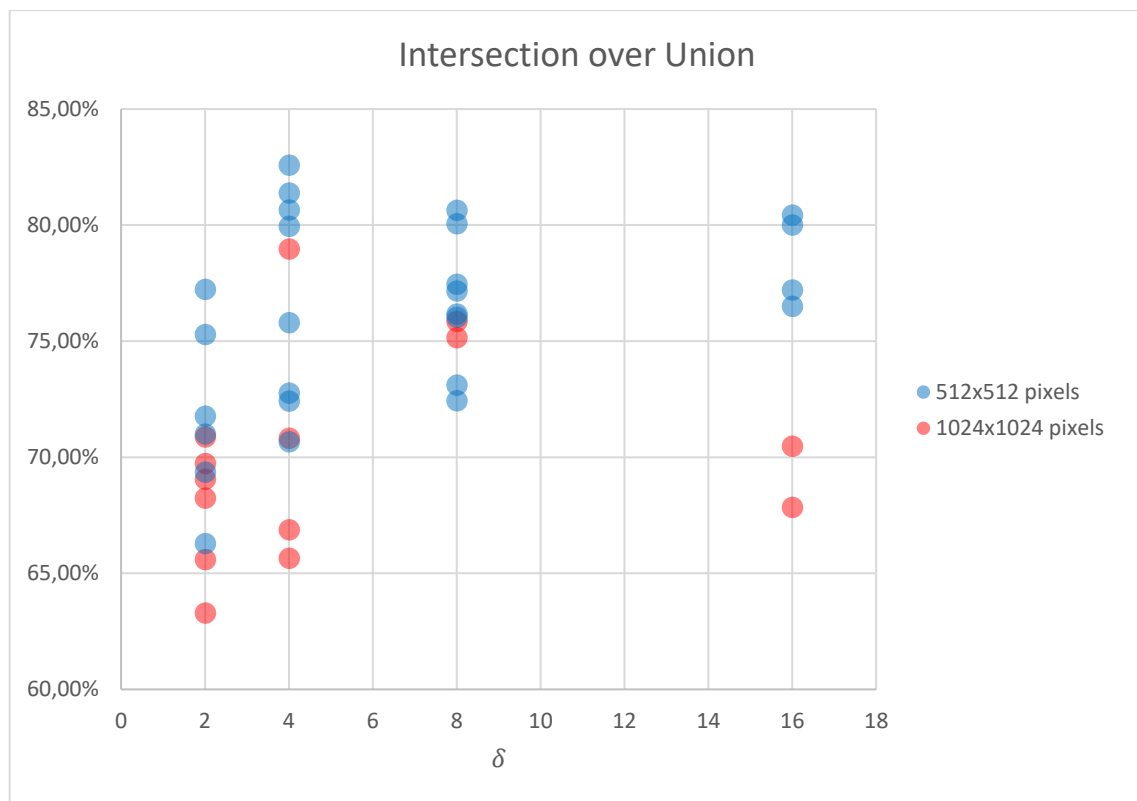
Fig. 5-3   Results of IoU for each model depth multiplication factor δ.

On Fig. 5-4 is graph which shows IoU for different model depth multiplication factor δ. This is for categorical cross entropy loss function and input/output image size of 512x512 pixels. Models with the lowest complexity show decrease of performance with higher batch sizes. Models with higher complexity show increasing performance with higher batch sizes. Model with model depth multiplication factor δ of 16 shows only two data points because of OOM error during training.
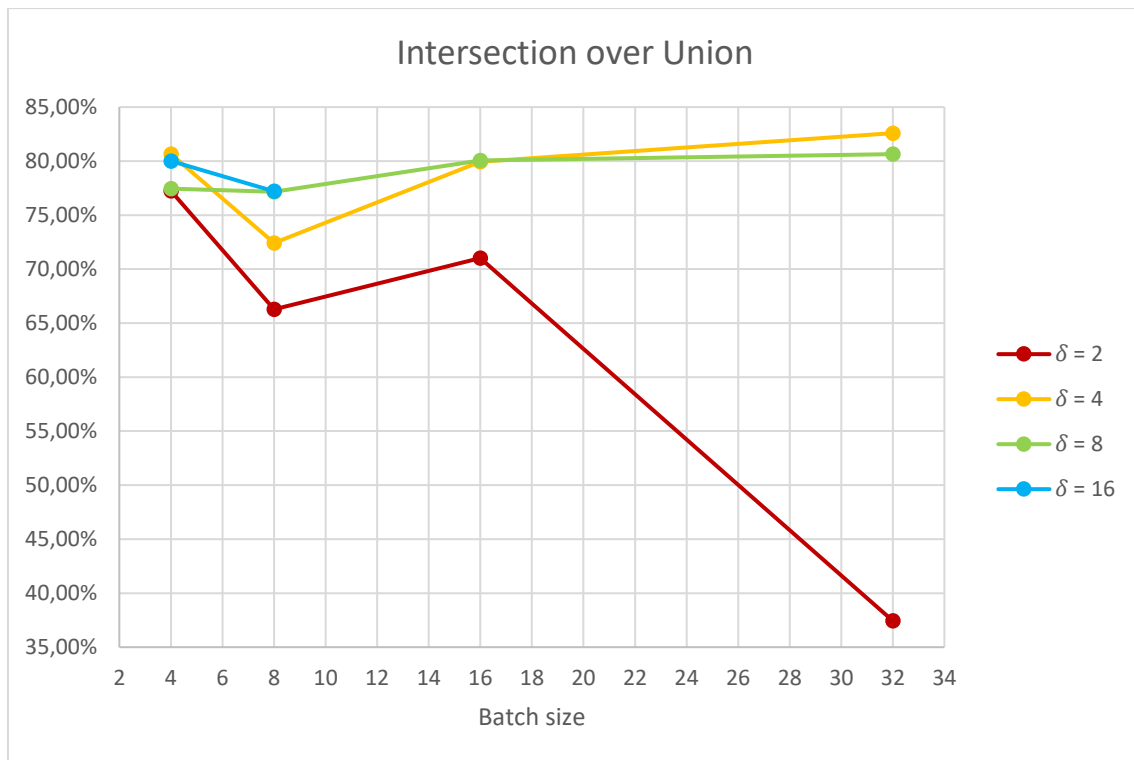
Fig. 5-4 Relation of IoU to the batch size for each multiplication factor δ. Performed for cross entropy loss and input image size of 512x512 pixels.

Graph on Fig. 5-5 shows relation of PA to the model depth multiplication factor δ for cross entropy loss, input/output image size of 512x512 and batch size of 4. PA increases between value 2 to 4 but then starts to decrease when model starts to be more complex and has more trainable parameters. Values of PA for Gas pores have the lowest values, so model has hard time to classify this class.
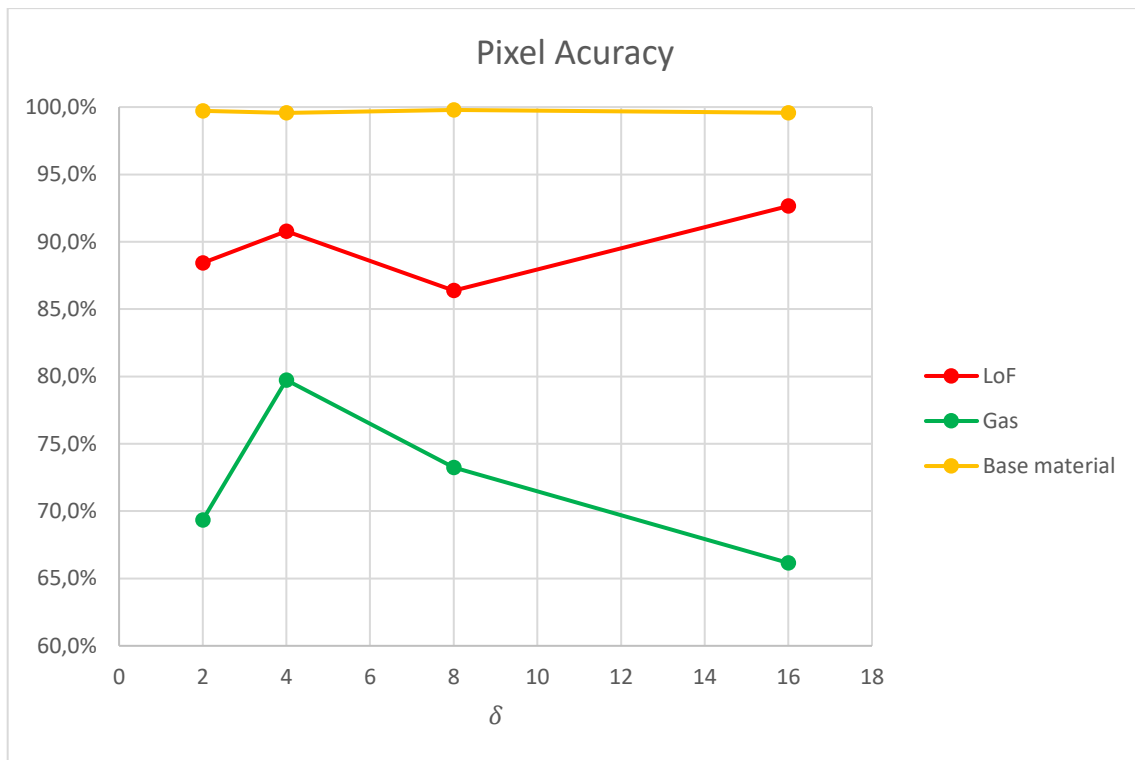
Fig. 5-5   PA for each model depth multiplication factor δ. Performed for cross entropy loss, input/output image size of 512x512 pixels and batch size of 4.

Graph on Fig. 5-6 shows values of PA relation to the batch size for cross entropy loss function, input/output image size of 512x512 and model depth multiplication factor δ of 4. The value of PA on gas pores class has again much lower values than lack of fusion pores or base material with highest value in batch size value of 4.
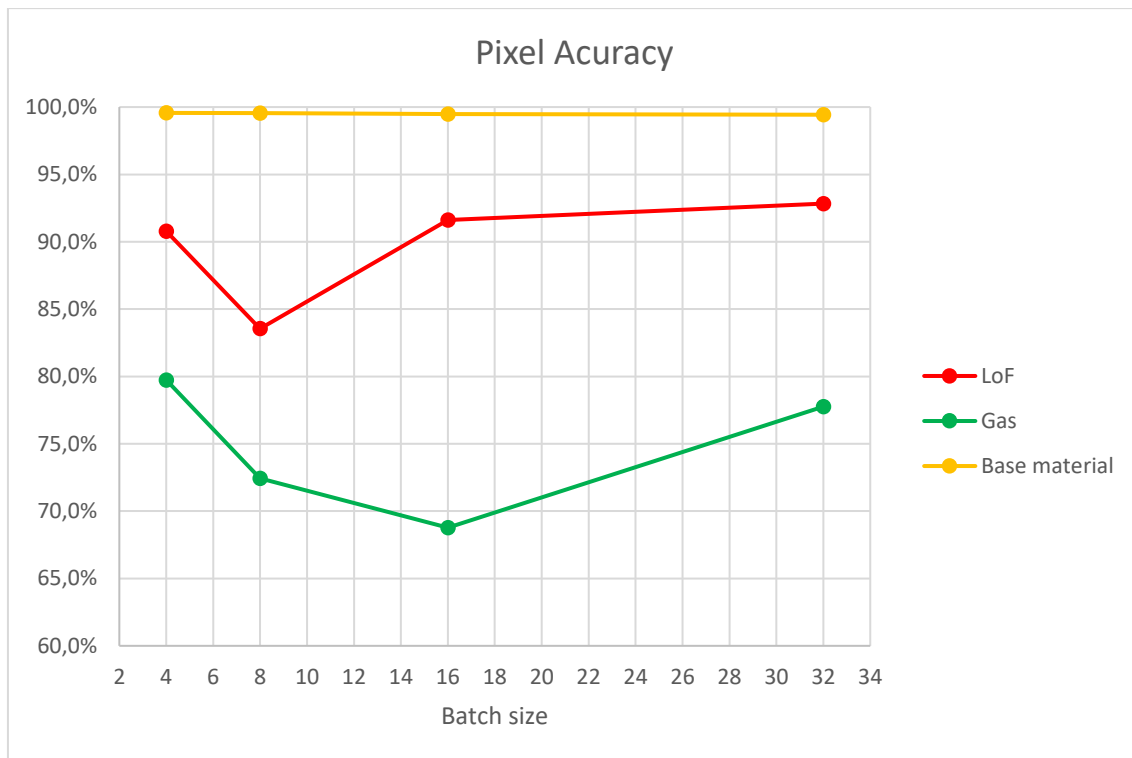
Fig. 5-6   PA for each batch size. Performed with cross entropy loss, input/output image size of 512x512 pixels and model depth multiplication factor δ of 4.

5.1.3   The best model evaluation

From all performed experiments I have picked model which was trained on categorical cross entropy loss function with model depth multiplication factor δ of 4, batch size of 4 and input/output image patch of 512x512 pixels. This model had one of the highest values of PA and IoU. It was picked also because it had reasonably low complexity due to low depth multiplication factor δ which correspond to 121 663 model parameters so it will not accommodate high amount of memory on the system during further implementation. Also input/output image size of 512x512 pixels is more suitable for further implementation because with this image size we can input to the network also images with lower sizes (the lowest size of 512x512 pixels). If we would like to segment images with higher sizes, we can simply crop the image to smaller 512x512 pixels sized patches.

The duration of training of this model was 50 minutes. On Fig. 5-7 it is shown how loss curve has developed during the training process. Curve shows loss calculated on the same data which was model trained on (train loss) and on validation data (validation loss). The training was stopped when validation loss was at its lowest point. That is the reason why I have saved only model which was in 1312th iteration. Validation loss has higher values due to calculation on data that model has not seen before compared to train data. Similarly, train PA, validation PA and IoU have been calculated during development of the training process. These are shown on Fig. 5-8 and Fig. 5-9.
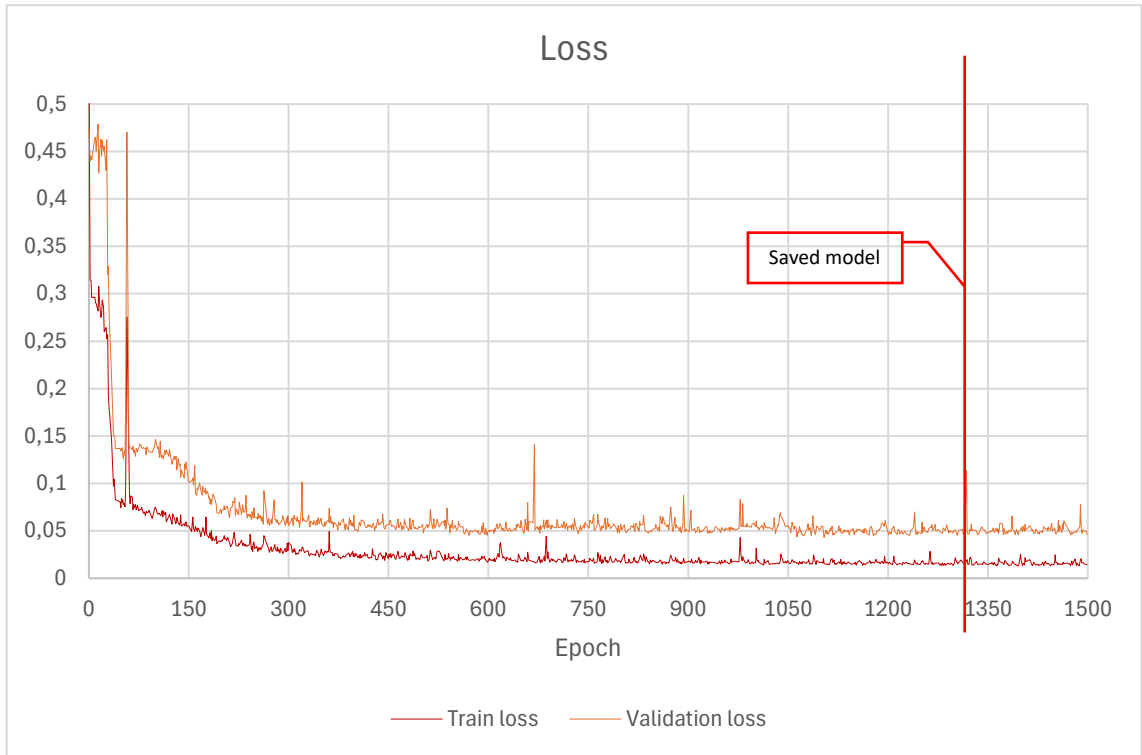
Fig. 5-7   Cross entropy train loss and validation loss development during training of picked U-net model.
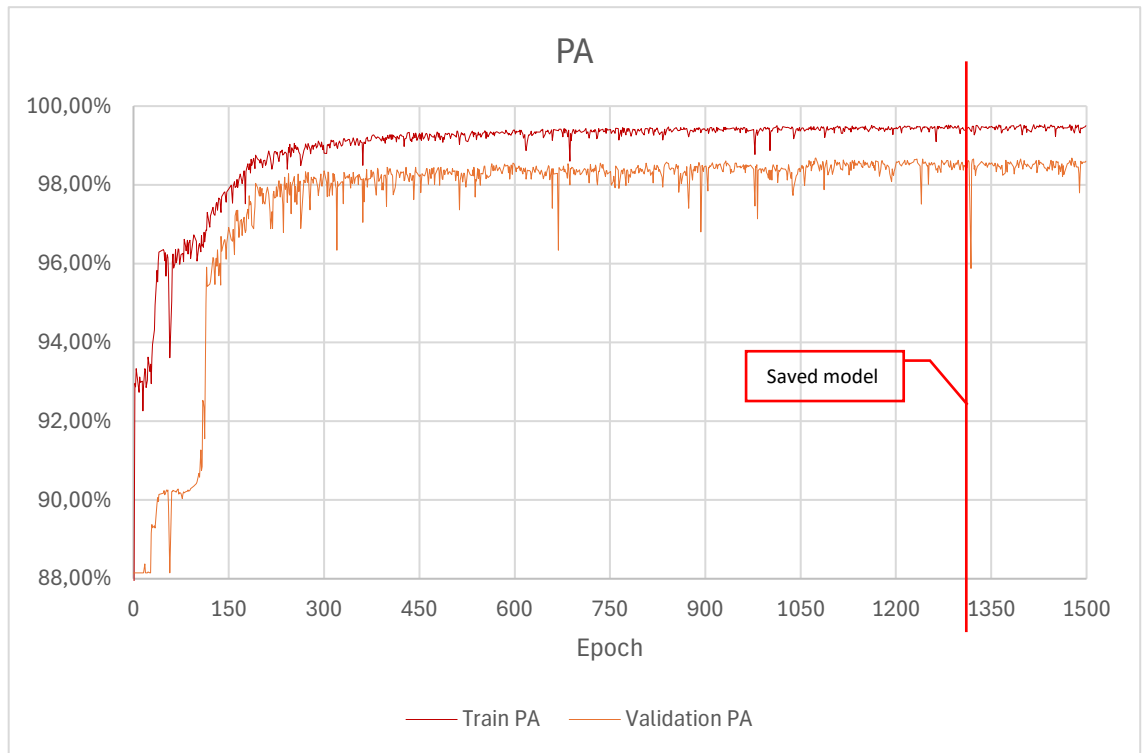
Fig. 5-8   Train PA and validation PA development during training of picked U-net model.
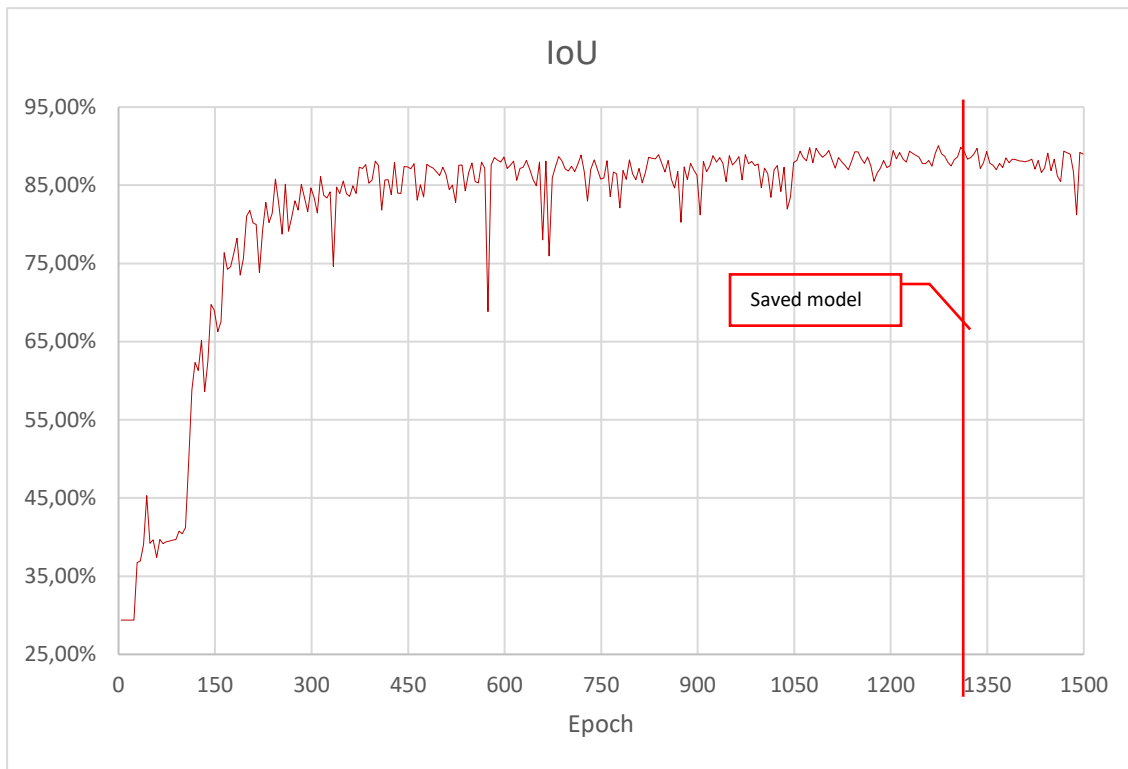
Fig. 5-9   IoU development during training of picked U-net model.

Examples of segmentation maps which were created with this model on test set of images are shown on Fig. 5-10. On Fig. 5-10a) model did not correctly predicted part of lack of fusion porosity. On Fig. 5-10b) model correctly predicted gas porosity but on the edge of one pore, model predicts lack of fusion porosity. On original image there is dried cleaning medium on this place. On Fig. 5-10c) there is some kind of fiber sticking out of pore which model predicts as lack of fusion porosity. The model predicts small parts inside of lack of fusion porosity as gas porosity.

| Original image | Predicted mask | Ground truth |

a)

b)

c)

d)

🔴 Lack of fusion 🟢 Gas porosity 🟡 Base material
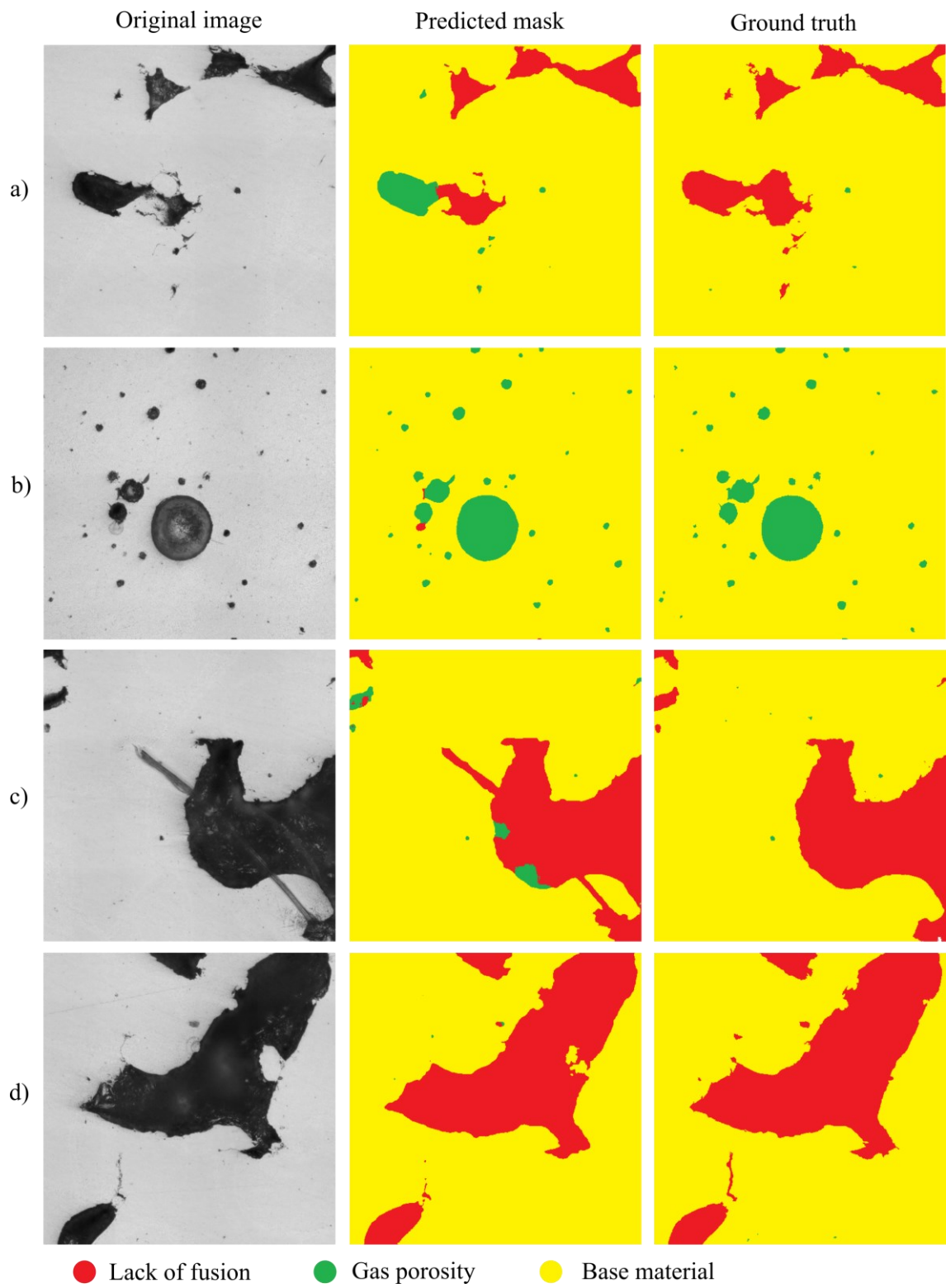
Fig. 5-10 AlSi10Mg, examples of predicted masks from test set, compared with original images and ground truth masks (each 512x512 pixels).

5.1.4 Testing on Cu99 and Ti6Al4V

For testing on Cu99 and Ti6Al4V, the same model which was trained on AlSi10Mg was used. On Tab. 5-2 there are resulting performance metrices for this testing. We can see from the table that model has no problem with predicting base material class due to high percentage of PA. Fig. 5-11 and Fig. 5-12 shows examples of prediction maps from selected images for Cu99 and Ti6Al4V.

For Cu99 on Fig. 5-11a) there are a lot of small lack of fusion strings which model is not able to recognize, also it wrongly predicts parts of lack of fusion pores, and it is replacing it with gas porosity class. On Fig. 5-11b) model also predicts parts of lack of fusion pores as gas porosity.

For Ti6Al4V on Fig. 5-12a) model predicts most of the parts of lack of fusion porosity well. On Fig. 5-12b) model again has a problem with gas porosity class.

Tab. 5-2   Performance metrices of model evaluated on Cu99 and Ti6Al4V

Material	PA (LoF)	PA (Gas)	PA (Base)	IoU (Mean)
Cu99	41,95%	97,85%	99,57%	51,27%
Ti6Al4V	94,52%	52,83%	99,00%	73,24%

Fig. 5-11   Cu99, examples of predicted masks, compared with original images and ground truth masks (each 512x512 pixels).

Fig. 5-12 Ti6Al4V, examples of predicted masks, compared with original images and ground truth masks (each 512x512 pixels).

5.1.5 Comparison with threshold based segmentation

For comparison of trained model, threshold-based image segmentation was used which is classical method used for this purpose at this moment. For finding the threshold value, Otsu's method was used. For comparison, I have also computed PA of prediction maps which were edited so there are not three classes on the image but only two classes, same as within threshold-based segmentation. Resulting PA compared with labeled ground truth are shown in Tab. 5-3. From the table it is observable that if we compare threshold-based segmentation with the model predictions, the AlSi10Mg model predicts defects and base material better. Examples of images which were segmented with threshold-based segmentation compared with model predictions for AlSi10Mg are shown on Fig. 5-13, for Cu99 on Fig. 5-14 and for Ti6Al4V on Fig. 5-15.

61

Tab. 5-3 PA of threshold-based segmentation and model predictions for two classes.

Material	Threshold		Model prediction	
	Defects	Base material	Defects	Base material
AlSi10Mg	93,67%	99,71%	94,24%	99,57%
Cu99	70,36%	99,30%	57,28%	99,57%
Ti6Al4V	95,51%	89,01%	96,38%	99,00%
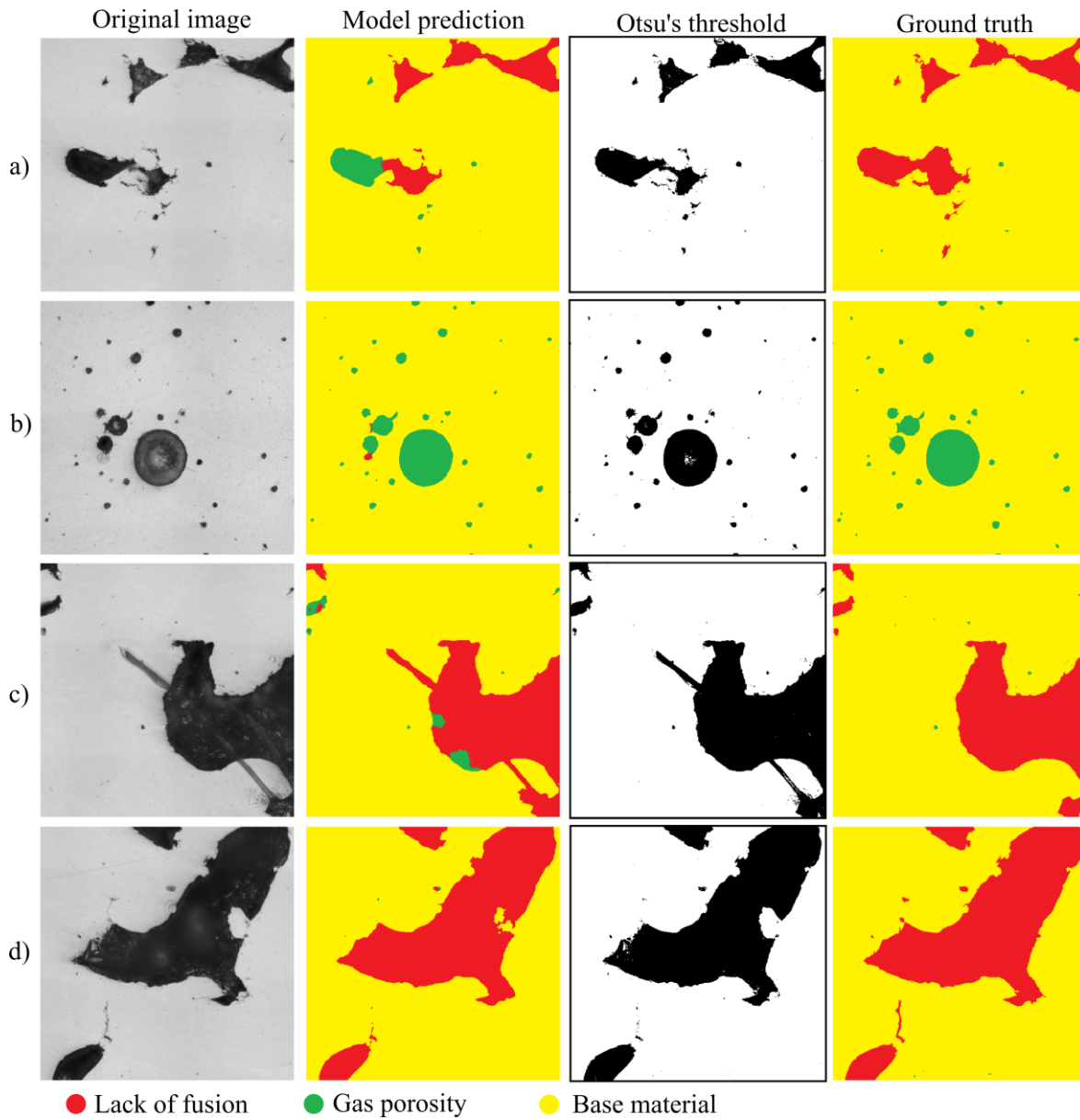
Fig. 5-13 AlSi10Mg, examples of predicted masks compared with original images and threshold-based segmentation (each 512x512 pixels).

| Original image | Model prediction | Otsu's threshold | Ground truth |

a)

b)

● Lack of fusion ● Gas porosity ● Base material
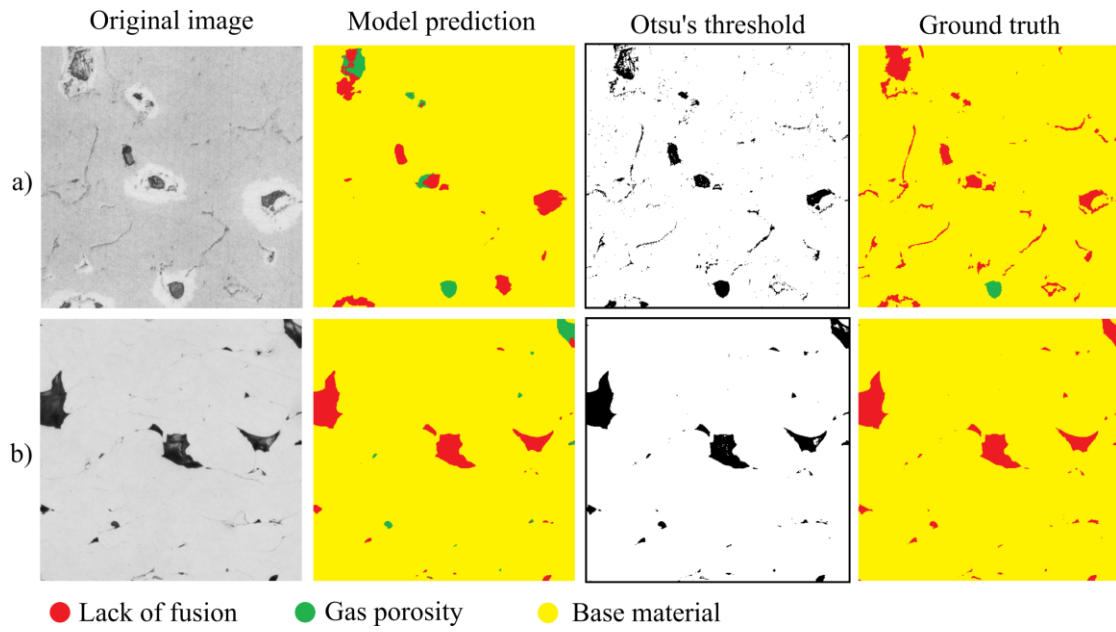
Fig. 5-14   Cu99, examples of predicted masks compared with original images and threshold-based segmentation (each 512x512 pixels).

| Original image | Model prediction | Otsu's threshold | Ground truth |

a)

b)

● Lack of fusion ● Gas porosity ● Base material
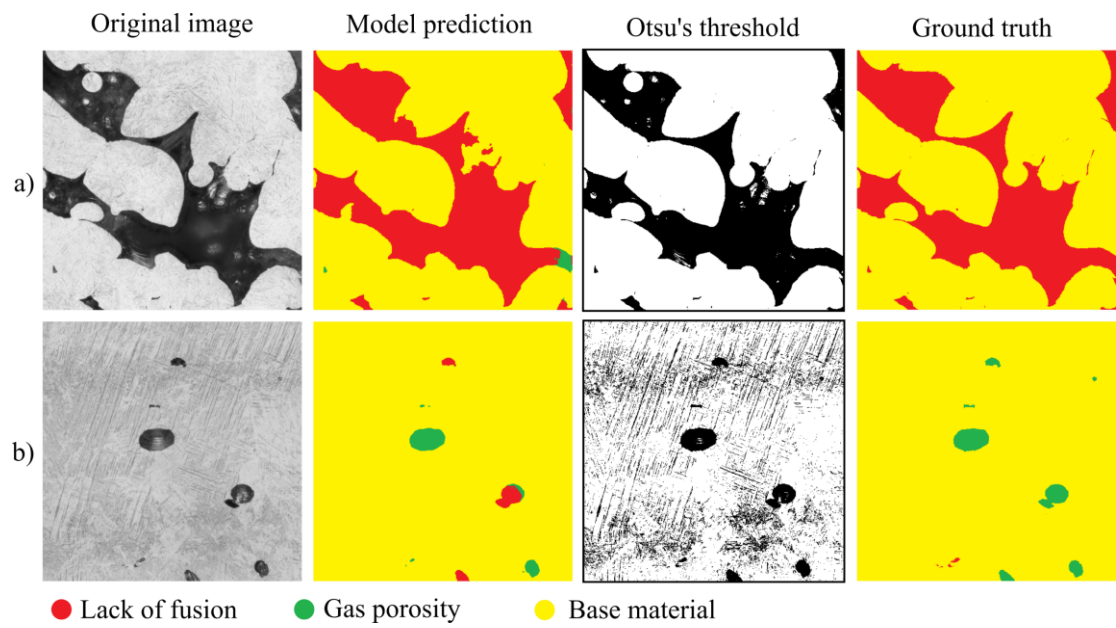
Fig. 5-15   Ti6Al4V, examples of predicted masks compared with original images and threshold-based segmentation (each 512x512 pixels).

5.2 Image super-resolution

Only initial experiments with image-super resolution neural network training were performed. On Fig. 5-16 and Fig. 5-17 there are graphs which show how performance metrics of SSIM and PSNR developed during the training process. These were calculated on validation data. Training was stopped after 1760 training iterations and it took 44 hours. At the end of training, the network was able to generate images which had 0,767 of SSIM and 34,13 dB PSNR on validation data.
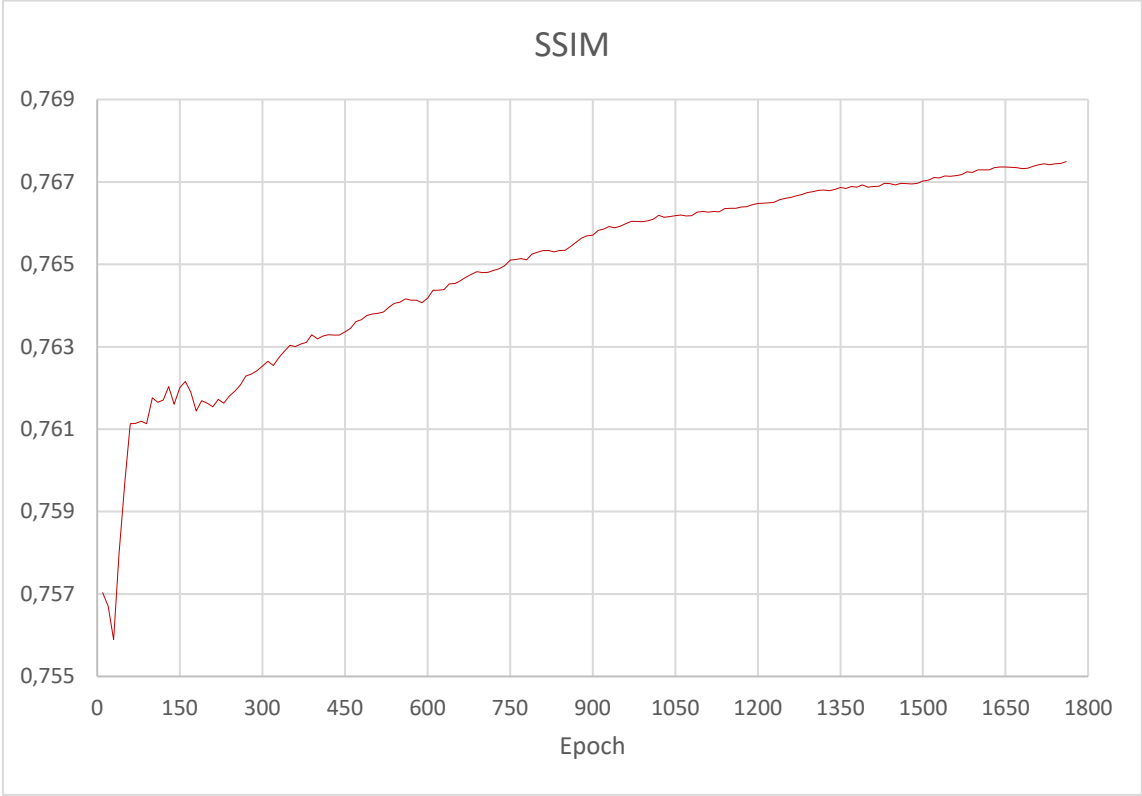
Fig. 5-16   SSIM development during training of picked SRGAN model.
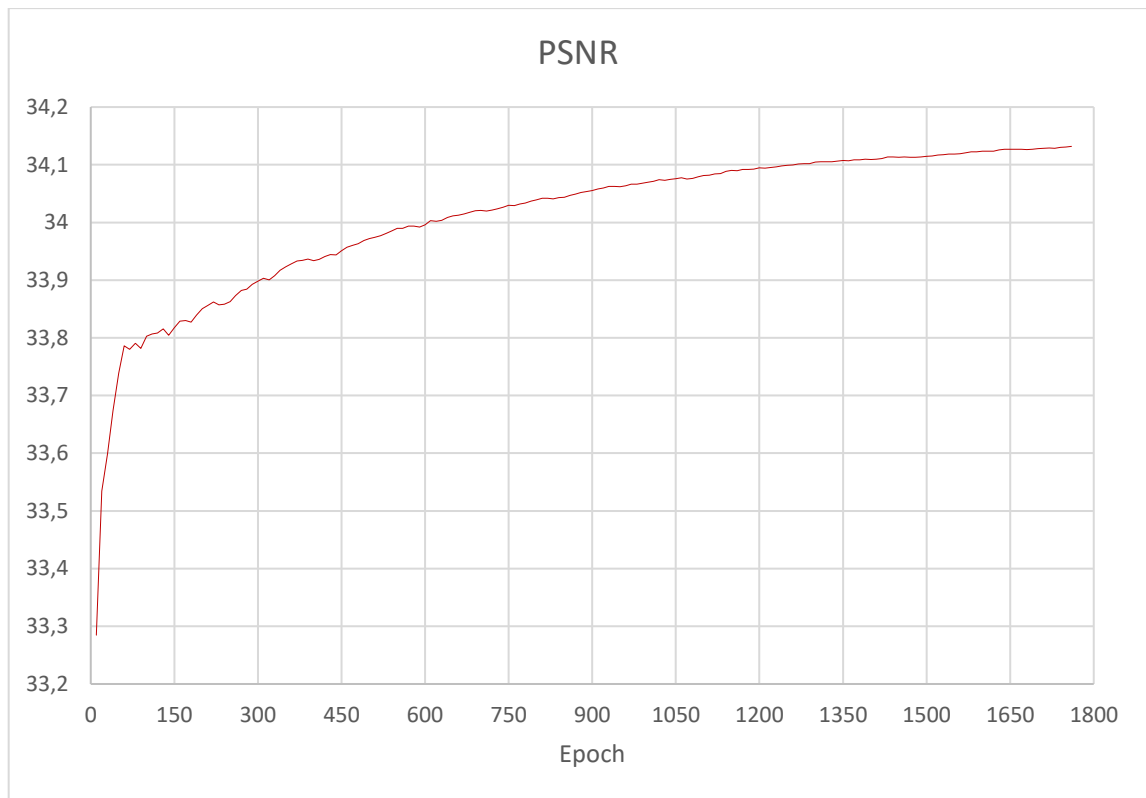
Fig. 5-17   PSNR development during training of picked SRGAN model.

Examples of images which were generated with trained SRGAN compared with original high resolution and low-resolution images are shown on are shown on Fig. 5-18. The network was able to smooth the edges of pores and bring back the details of the image.

Generated image   Original HR   Original LR
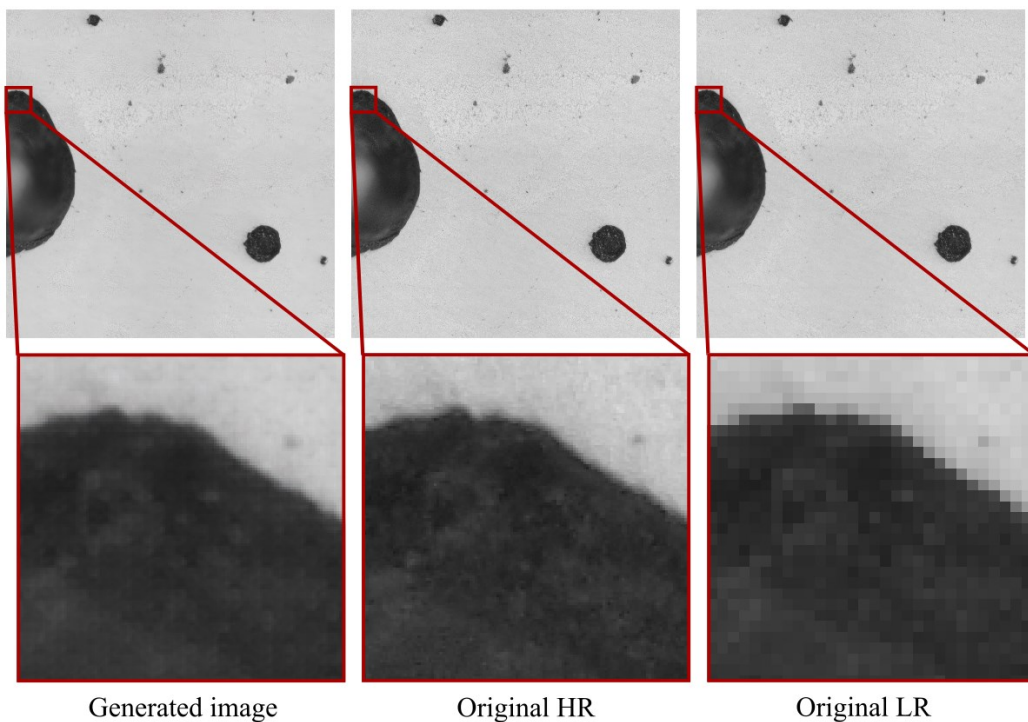
Fig. 5-18   AlSi10Mg, example of generated image compared with original HR and LR image.

5.3 Implementation

Best trained U-net model further explained in 5.1.3 was implemented in custom solution with simple graphical user interface (GUI). Executable files have been created and tested on Linux operating system. Implementation can also be executed as a script. Simple interface is shown on Fig. 5-19. User must choose input folder, which contain images with materials cross-sections with defects. Images must be in .tif format and created with 250x magnification lens with Keyence microscope mentioned in 4.2.1. After execution of script, percentual ratio of each defect and base material will be shown to the user on the interface and segmentation maps with corresponding original image patches will be saved in save folder directory (Fig. 5-20).

If the user wants to execute implementation as a python script, required libraries which need to be installed are included in directory in enviroment.yml file. It is recommended to use Anaconda environment for installing required libraries straight from included environment file [30].
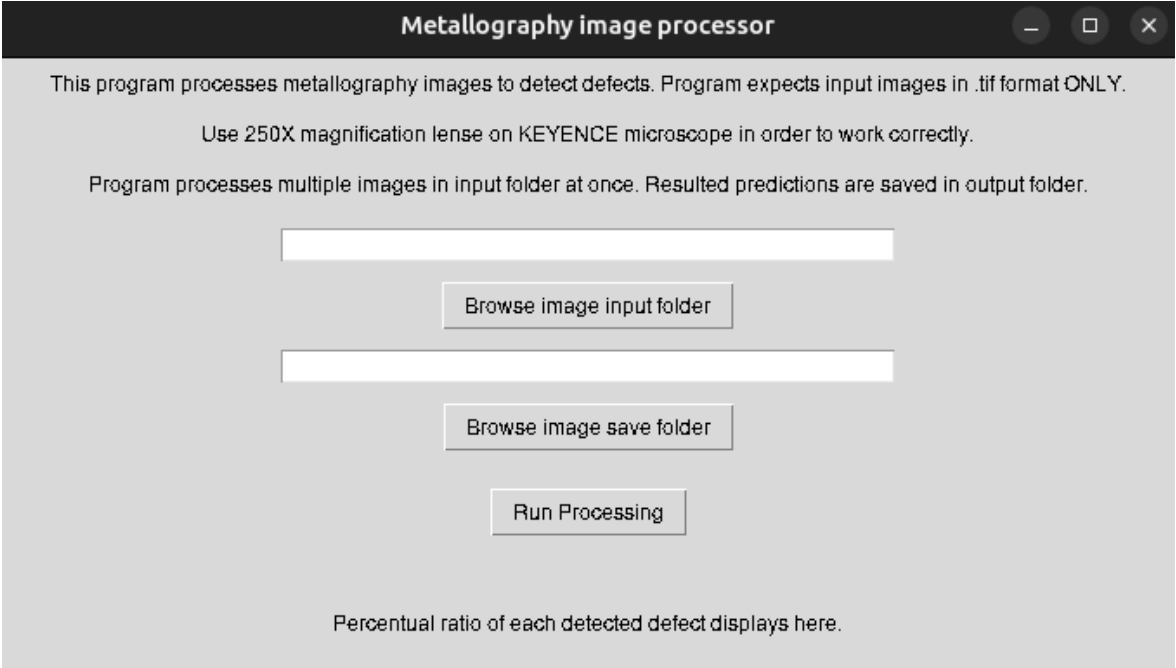
Fig. 5-19 Simple GUI for image processing implementation
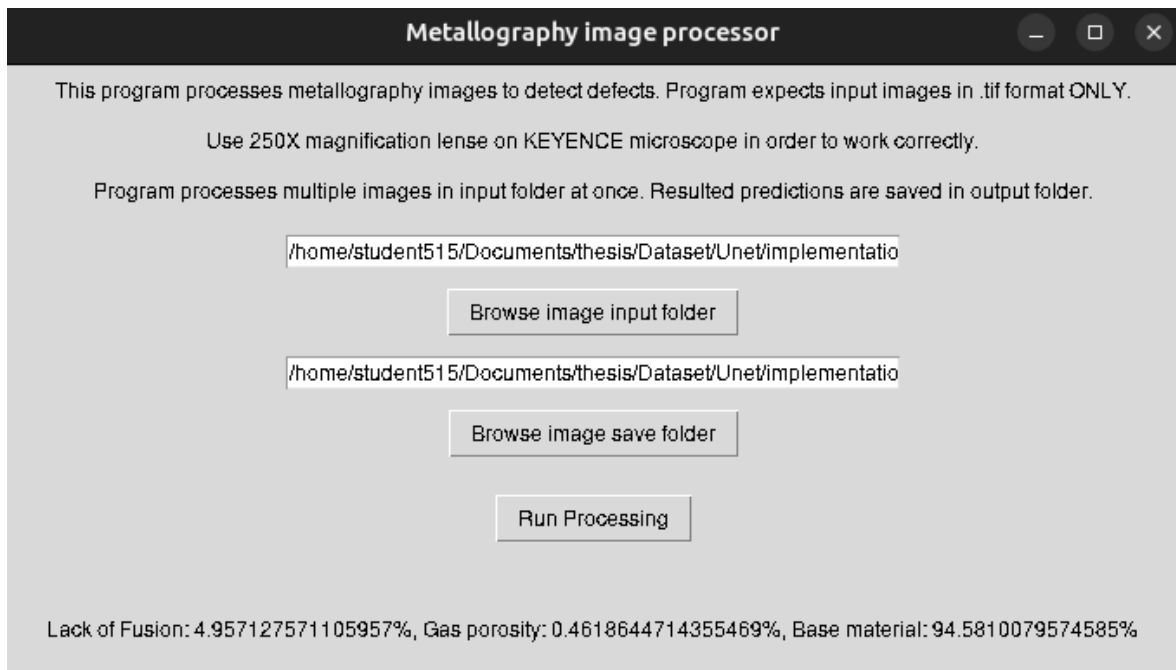
Fig. 5-20   Example of percentual results on GUI.

6　DISCUSSION

This work was focusing on categorizing the defects of metallographic cross sections to classes and finding and applying state of the art image processing methods for image evaluation, also by the goal of improving accuracy of the results of image segmentation. It has been found that these methods are really sophisticated and strongly rely on the quality of input training data, also on finding the best training setup parameters for the training process.

6.1　Interpretation of results

6.1.1　Dataset

During data labeling process, it is important to have professionals in the field creating the labels because if there is some uncertainty during labeling process, this will be mirrored in neural-network performance, and it will only perform as good as operator which was creating dataset. I also think that labels in dataset that I have created for AlSi10Mg were not perfect because during labeling process, I was also sometimes unsure if corresponding defect correspond to lack of fusion porosity or gas porosity.

This is also the reason why it might be beneficial to explore unsupervised learning methods for this task in the future, in which models can learn not only from labeled data but also during training by itself from unlabeled data. This was also stated by Luengo J et. al. [2].

6.1.2　Semantic segmentation

From the experiments addressed in 5.1.2 it is clear that used computer setup is not powerful enough and I could not perform all experiments because of lack of memory in graphics card. At the same time, it is clear that model still can perform sufficiently well with smaller image sizes (HxW) so I was still able to create model with sufficient performance for single material type and relatively small dataset.

With the best model I have managed to get performance of IoU of 80,66% with per class accuracy of 90,8% for Lack of Fusion porosity, 79,7% for Gas porosity and 99,6% for base material. These values are comparable with state-of-the-art mentioned in 2.2.2 (Tab. 2-5). To further improve the model, it would be sufficient to create a more reliable dataset by experts in the field. I also think that if we would like to train the model on more complex dataset with all three materials, we would need to use deeper model with more trainable parameters to understand all three materials in one complex model. Also, it would be beneficial to further experiment with different loss functions.

For my use case and model trained only on AlSi10Mg I have found that lower neural network input image sizes HxW had higher values of IoU metric for model with batch size of 4. I think that if I would like to use a higher image size of 1024x1024, a higher batch size would also need to be used during the training process but this needs to be performed on more powerful computing setup.

6.1.3 Segmentation of unseen data

It has been shown that the model trained on AlSi10Mg was not able to perform correct segmentation on Cu99 since the structures on these materials have different shapes. On Ti6Al4V, the model was able to perform accurate segmentation of Lack of Fusion porosity, because the shape of this defect is like the ones on AlSi10Mg. That is the reason why we can still use this model to predict on different material which the model was not trained on, but defect need to have similar shape. If there are some structures which do not look like the structures on AlSi10Mg, the model will perform poorly. At the same time, we can see that model predicts base material with PA of 99,57% for Cu99 and PA of 99,00% for Ti6Al4V (Tab. 5-2). That is because the structure of base material in gray scale looks the same cross all tested materials. Under this condition, we can use this model for detecting the percentual ratio of base material class.

6.1.4 Comparison with threshold-based segmentation.

From results on Tab. 5-3 it is observed that neural network model trained on AlSi10Mg can predict defects and base material comparable good as threshold base segmentation. Base material class was predicted with the same or with Ti6Al4V with even better accuracy than with threshold-based segmentation. This is because threshold-based segmentation identified a lot of parts of base material as defects even if it was base material as we can see on Fig. 5-15b). It can be stated that segmentation results from proposed model are as good as threshold-based segmentation, with the additional advantage of categorizing the defects into classes.

6.1.5 Image super-resolution

After initial experiments with image super resolution model SRGAN, performance of 0,767 of SSIM and 34,13 dB PSNR was achieved. These values are similar as values mentioned in previous works mentioned in 2.2.3 (Tab. 2-6).

After training of semantic segmentation model U-net, I have realized that my computer setup is not sufficient to process full sized images. That is the reason why I had to down scale images for training of semantic segmentation and I was not able to connect image super resolution model with semantic segmentation model.

6.2 Hypothesis verification

Hypothesis 1.: If image segmentation neural network model will be trained only on material AlSi10Mg, it will still be able to perform accurate segmentation of defects on material Cu99 and Ti6Al4V because of defects shape is simar cross all material types.

The results indicate that the proposed approach using a U-net neural network can achieve segmentation accuracy comparable to the previously used threshold-based segmentation method. However, a condition must be met: the shape of the defects must be the same as those that form on AlSi10Mg. Therefore, the hypothesis was rejected. To further study this, it would be necessary to train the network with higher amount of data with diverse spectrum of materials which would contain all the defect types.

Hypothesis 2.: If an image up-scaling algorithm based on GAN is used before image segmentation of metallographic cross sections, then the results of measured performance metrices will have higher value than the results of stand-alone segmentation.

Because of lack of computational power, I was not able to approve or disprove this hypothesis. With the usage of higher resolution, the defects on the image appear larger so also higher input/output image sizes are required for neural network in order to understand the context of the defects and its shapes. With this, the network becomes larger and accommodates a lot of memory, so it is not possible to be trained on regular computer setup. Thus, this hypothesis remains open and needs to be further studied.

7   CONCLUSION

Method for semantic segmentation based on deep neural networks for categorizing the defects on the images of 3D printed metallographic cross-sections and deeper understanding of defects is proposed. The method uses U-net neural network architecture trained on custom dataset of singe material AlSi10Mg. This dataset is relatively small but was artificially augmented during the training process.

With this method, the best obtained model had only 121 663 trainable parameters and achieved a performance of 80.66% IoU on test data that the model had not seen before, which is comparable to the state-of-the-art. With testing of this model on different material types (Cu99 and Ti6Al4V) model sometimes struggles with detection of required defect types due to defect shapes unsimilarities. Method is comparable with threshold-based segmentation with additional categorizing of defects to the classes. With threshold-based segmentation, it is crucial how well the specimens are prepared so there is not a lot of dark areas like with Ti6Al4V. The advantage of the proposed method is that it can be applied and still perform well on specimens which contain these dark areas, which with threshold-based segmentation would not be predicted correctly. Performed experiments with SRGAN generative neural network show potential in further generating accurate artificial data to enhance performance of image processing tasks.

The limitation of training the neural-network models is still computing power. This training is possible to be done on standard hardware setup only to a limited extent. For further experimentation also with models with higher number of trainable parameters and higher amount of training data, it is suitable to use computationally oriented graphics cards setup with high memory and computational power. The limitation of the proposed model is that it performs well only on dataset which it was trained on (AlSi10Mg).

To further enhance this method to be able to perform more robust and accurately, it is suitable to create dataset, which contains a broader spectrum of defects and materials. Additionally, the quality and noise of the labels are crucial for the performance of segmentation. Therefore, it is recommended that the dataset be created by professionals in the field. To improve the implementation and accuracy of resulting segmentation, it is also possible to combine multiple image post-processing techniques together with neural network model segmentation. Segmentation task could be further extended from 2D image data to 3D voxel mesh segmentation for segmenting 3D defect shapes from computational tomography meshes or from point clouds scans.

Further neural networks could be used in other metrology tasks in general for processing 2D or 3D data and finding relations between different variables or functions. In general, neural networks can be used for estimation and finding deep correlations between input and output data not only in the field of 3D printing and metallography.

8 REFERENCES

[1] LECUN, Yann, Yoshua BENGIO and Geoffrey HINTON. Deep learning. *Nature* [online]. 2015, **521**(7553), 436–444 [Accessed 2021-11-13]. ISSN 14764687. Available at: doi:10.1038/NATURE14539

[2] LUENGO, Julián, Raúl MORENO, Iván SEVILLANO, David CHARTE, Adrián PELÁEZ-VEGAS, Marta FERNÁNDEZ-MORENO, Pablo MESEJO and Francisco HERRERA. A tutorial on the segmentation of metallographic images: Taxonomy, new MetalDAM dataset, deep learning-based ensemble model, experimental analysis and challenges. *Information Fusion* [online]. 2022, **78**, 232–253 [Accessed 2022-01-12]. ISSN 15662535. Available at: doi:10.1016/J.INFFUS.2021.09.018

[3] ALZUBAIDI, Laith, Jinglan ZHANG, Amjad J. HUMAIDI, Ayad AL-DUJAILI, Ye DUAN, Omran AL-SHAMMA, J. SANTAMARÍA, Mohammed A. FADHEL, Muthana AL-AMIDIE and Laith FARHAN. Review of deep learning: concepts, CNN architectures, challenges, applications, future directions. *Journal of Big Data* [online]. 2021, **8**(1) [Accessed 2022-03-15]. ISSN 21961115. Available at: doi:10.1186/S40537-021-00444-8

[4] YUAN, Xiaohui, Jianfang SHI and Lichuan GU. A review of deep learning methods for semantic segmentation of remote sensing imagery. *Expert Systems with Applications* [online]. 2021, **169** [Accessed 2022-03-15]. ISSN 09574174. Available at: doi:10.1016/J.ESWA.2020.114417

[5] CHEN, Honggang, Xiaohai HE, Linbo QING, Yuanyuan WU, Chao REN, Ray E. SHERIFF and Ce ZHU. Real-world single image super-resolution: A brief review. *Information Fusion* [online]. 2022, **79**, 124–145 [Accessed 2022-04-25]. ISSN 1566-2535. Available at: doi:10.1016/J.INFFUS.2021.09.005

[6] LEDIG, Christian, Lucas THEIS, Ferenc HUSZÁR, Jose CABALLERO, Andrew CUNNINGHAM, Alejandro ACOSTA, Andrew AITKEN, Alykhan TEJANI, Johannes TOTZ, Zehan WANG and Wenzhe SHI. Photo-realistic single image super-resolution using a generative adversarial network. *Proceedings - 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017* [online]. 2017, **2017-January**, 105–114 [Accessed 2022-03-15]. Available at: doi:10.1109/CVPR.2017.19

[7] WANG, Xintao, Ke YU, Shixiang WU, Jinjin GU, Yihao LIU, Chao DONG, Yu QIAO and Chen Change LOY. ESRGAN: Enhanced super-resolution generative adversarial networks. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* [online]. 2019, **11133 LNCS**, 63–79 [Accessed 2021-12-10]. ISSN 16113349. Available at: doi:10.1007/978-3-030-11021-5_5

[8]   DECOST, Brian L, Bo LEI, Toby FRANCIS and Elizabeth A HOLM. High throughput quantitative metallography for complex microstructures using deep learning: A case study in ultrahigh carbon steel. *MICROSCOPY AND MICROANALYSIS* [online]. 2019 [Accessed 2021-10-08]. Available at: doi:10.1017/S1431927618015635

[9]   DEBROY, T., T. MUKHERJEE, H. L. WEI, J. W. ELMER and J. O. MILEWSKI. Metallurgy, mechanistic models and machine learning in metal printing. *Nature Reviews Materials* [online]. 2021, **6**(1), 48–68 [Accessed 2022-05-01]. ISSN 20588437. Available at: doi:10.1038/S41578-020-00236-1

[10]   MURR, L. E. A Metallographic Review of 3D Printing/Additive Manufacturing of Metal and Alloy Products and Components. *Metallography, Microstructure, and Analysis* [online]. 2018, **7**(2), 103–132 [Accessed 2022-03-14]. ISSN 21929270. Available at: doi:10.1007/S13632-018-0433-6

[11]   ROBERTS, Graham, Simon Y. HAILE, Rajat SAINJU, Danny J. EDWARDS, Brian HUTCHINSON and Yuanyuan ZHU. Deep Learning for Semantic Segmentation of Defects in Advanced STEM Images of Steels. *Scientific Reports* [online]. 2019, **9**(1). ISSN 20452322. Available at: doi:10.1038/s41598-019-49105-0

[12]   DE HAAN, Kevin, Zachary S. BALLARD, Yair RIVENSON, Yichen WU and Aydogan OZCAN. Resolution enhancement in scanning electron microscopy using deep learning. *Scientific Reports 2019 9:1* [online]. 2019, **9**(1), 1–7 [Accessed 2021-10-08]. ISSN 2045-2322. Available at: doi:10.1038/s41598-019-48444-2

[13]   NOWROTH, Christian, Tiansheng GU, Jan GRAJCZAK, Sarah NOTHDURFT, Jens TWIEFEL, Jörg HERMSDORF, Stefan KAIERLE and Jörg WALLASCHEK. Deep Learning-Based Weld Contour and Defect Detection from Micrographs of Laser Beam Welded Semi-Finished Products. *Applied Sciences (Switzerland)* [online]. 2022, **12**(9). ISSN 20763417. Available at: doi:10.3390/app12094645

[14]   SHEN, Chunguang, Chenchong WANG, Minghao HUANG, Ning XU, Sybrand VAN DER ZWAAG and Wei XU. A generic high-throughput microstructure classification and quantification method for regular SEM images of complex steel microstructures combining EBSD labeling and deep learning. *Journal of Materials Science & Technology* [online]. 2021, **93**, 191–204 [Accessed 2022-03-14]. ISSN 1005-0302. Available at: doi:10.1016/J.JMST.2021.04.009

[15]   SHEN, Chunguang, Xiaolu WEI, Chenchong WANG and Wei XU. A deep learning method for extensible microstructural quantification of DP steel enhanced by physical metallurgy-guided data augmentation. *Materials Characterization* [online]. 2021, **180** [Accessed 2022-03-14]. ISSN 10445803. Available at: doi:10.1016/J.MATCHAR.2021.111392

[16] LI, Wenhao, Ding HE, Yongqiang LIU, Fenghe WANG and Fengliang HUANG. Super-resolution reconstruction, recognition, and evaluation of laser confocal images of hyperaccumulator Solanum nigrum endocytosis vesicles based on deep learning: Comparative study of SRGAN and SRResNet. *Frontiers in Plant Science* [online]. 2023, **14**. ISSN 1664462X. Available at: doi:10.3389/fpls.2023.1146485

[17] BANSAL, Aayush, Xinlei CHEN, Bryan RUSSELL, Abhinav GUPTA and Deva RAMANAN. PixelNet: Representation of the pixels, by the pixels, and for the pixels [online]. 2017 [vid. 2022-04-27]. Available at: doi:10.48550/arxiv.1702.06506

[18] RONNEBERGER, Olaf, Philipp FISCHER and Thomas BROX. U-Net: Convolutional Networks for Biomedical Image Segmentation. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* [online]. 2015, **9351**, 234–241 [Accessed 2022-04-27]. ISSN 16113349. Available at: doi:10.1007/978-3-319-24574-4_28

[19] BRIAN L. DECOST ET AL. *UHCSDB (UltraHigh Carbon Steel micro-graph DataBase): tools for exploring large heterogeneous microstruc-ture datasets* [online]. B.m.: Integrating Materials and Manufacturing Innovation. 2017 [Accessed 2022-04-27]. Available at: http://uhcsdb.materials.cmu.edu

[20] NOLDEN, Marco, Sascha ZELZER, Alexander SEITEL, Diana WALD, Michael MÜLLER, Alfred M. FRANZ, Daniel MALEIKE, Markus FANGERAU, Matthias BAUMHAUER, Lena MAIER-HEIN, Klaus H. MAIER-HEIN, Hans Peter MEINZER and Ivo WOLF. The medical imaging interaction toolkit: Challenges and advances: 10 years of open-source development. *International Journal of Computer Assisted Radiology and Surgery* [online]. 2013, **8**(4), 607–620 [Accessed 2022-04-27]. ISSN 18616429. Available at: doi:10.1007/S11548-013-0840-8/FIGURES/10

[21] SINGLA, Khushboo, Rajoo PANDEY and Umesh GHANEKAR. A review on Single Image Super Resolution techniques using generative adversarial network. *Optik* [online]. 2022, **266** [Accessed 2024-03-14]. ISSN 00304026. Available at: doi:10.1016/J.IJLEO.2022.169607

[22] CHEN, Dali, Dinghao GUO, Shixin LIU and Fang LIU. Microstructure instance segmentation from aluminum alloy metallographic image using different loss functions. *Symmetry* [online]. 2020, **12**(4). ISSN 20738994. Available at: doi:10.3390/SYM12040639

[23] HORÉ, Alain and Djemel ZIOU. Image quality metrics: PSNR vs. SSIM. In: *Proceedings - International Conference on Pattern Recognition* [online]. 2010, s. 2366–2369. ISBN 9780769541099. Available at: doi:10.1109/ICPR.2010.579

[24] JADON, Shruti. A survey of loss functions for semantic segmentation. *IEEE Conference on Computational Intelligence in Bioinformatics and Computational Biology (CIBCB)* [online]. 2020 [Accessed 2024-05-01]. Available at: https://github.com/shruti-jadon/

[25] LIN, Tsung Yi, Priya GOYAL, Ross GIRSHICK, Kaiming HE and Piotr DOLLAR. Focal Loss for Dense Object Detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence* [online]. 2017, **42**(2), 318–327 [Accessed 2024-05-07]. ISSN 19393539. Available at: doi:10.1109/TPAMI.2018.2858826

[26] FRIZZA, Tristan, Donald G. DANSEREAU, Nagita Mehr SERESHT and Michael BEWLEY. Semantically accurate super-resolution Generative Adversarial Networks. *Computer Vision and Image Understanding* [online]. 2022, **221**. ISSN 1090235X. Available at: doi:10.1016/j.cviu.2022.103464

[27] Digital Microscope VHX-6000. *KEYENCE International Belgium* [online]. [Accessed 2023-11-05]. Available at: https://www.keyence.co.uk/products/microscope/digital-microscope/vhx-6000/models/vhx-6000/

[28] CUDA GPUs - Compute capability. *NVIDIA Developer* [online]. [Accessed 2023-11-05]. Available at: https://developer.nvidia.com/cuda-gpus

[29] ABADI, Martín, Ashish AGARWAL, Paul BARHAM, Eugene BREVDO, Zhifeng CHEN, Craig CITRO, Greg S CORRADO, Andy DAVIS, Jeffrey DEAN, Matthieu DEVIN, Sanjay GHEMAWAT, Ian GOODFELLOW, Andrew HARP, Geoffrey IRVING, Michael ISARD, Yangqing JIA, Rafal JOZEFOWICZ, Lukasz KAISER, Manjunath KUDLUR, Josh LEVENBERG, Dan MANÉ, Rajat MONGA, Sherry MOORE, Derek MURRAY, Chris OLAH, Mike SCHUSTER, Jonathon SHLENS, Benoit STEINER, Ilya SUTSKEVER, Kunal TALWAR, Paul TUCKER, Vincent VANHOUCKE, Vijay VASUDEVAN, Fernanda VIÉGAS, Oriol VINYALS, Pete WARDEN, Martin WATTENBERG, Martin WICKE, Yuan YU, Xiaoqiang ZHENG and Google RESEARCH. TensorFlow: Large-Scale Machine Learning on Heterogeneous Distributed Systems [online]. undated [Accessed 2024-05-04]. Available at: www.tensorflow.org.

[30] CONTINUUM ANALYTICS. *Anaconda. Version 4.9.2* [online]. 2020 [Accessed 2024-05-23]. Available at: https://www.anaconda.com

[31] OTSU, Nobuyuki. THRESHOLD SELECTION METHOD FROM GRAY-LEVEL HISTOGRAMS. *IEEE Trans Syst Man Cybern* [online]. 1979, **SMC-9**(1), 62–66 [Accessed 2024-05-16]. ISSN 00189472. Available at: doi:10.1109/TSMC.1979.4310076

[32]   KINGMA, Diederik P. and Jimmy Lei BA. Adam: A Method for Stochastic Optimization. *3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings* [online]. 2014 [Accessed 2024-05-07]. Available at: https://arxiv.org/abs/1412.6980v9

9   LIST OF USED ABBREVIATIONS, SYMBOLS AND QUANTITIES

9.1   List of abbreviations

DP	Dual phase
EBSD	Electron backscattered diffraction
ESRGAN	Enhanced super-resolution generative adversarial network
GAN	Generative adversarial network
Gas	Gas porosity
GPU	Graphics processing unit
GUI	Graphical user interface
HR	High resolution
IoU	Region intersection over union
LoF	Lack of fusion porosity
LR	Low resolution
MITK	Medical imaging interaction toolkit
MSE	Mean square error
OOM	Out of memory
PA	Pixel accuracy
PSNR	Peak signal-to-noise ratio
PRISMA	Preferred reporting items for systematic reviews and meta-analyses
Q&P	Quenching and partitioning
ReLu	Reflected linear unit
RGB	Red, green, blue
SEM	Scanning electron microscope
SSIM	Structure similarity index

SRGAN	Super-resolution generative adversarial network
UHCS	Ultra high carbon steel
2D	Two dimensional
3D	Three dimensional

9.2   List of symbols

C_1, C_2, C_3	Positive constant
c	Contrast comparison function
δ	Depth multiplication factor
f	Ground truth image
g	Test image
W	Width of image
H	Height of image
l	Luminance comparison function
s	Structure comparison function
σ_f	Standard deviation of ground truth image
σ_g	Standard deviation of test image
μ_f	Mean luminance of ground truth image
μ_g	Mean luminance of test image

10 LIST OF FIGURES AND GRAPHS

80

11 LIST OF TABES

12 LIST OF ATTACHMENTS

executable_dataset_scripts.zip