Article

Efficient Phase Segmentation of Light-Optical Microscopy Images of Highly Complex Microstructures Using a Correlative Approach in Combination with Deep Learning Techniques

Björn-Ivo Bachmann [1,2,*] , Martin Müller [1,2], Marie Stiefel [1,2] , Dominik Britz [1,2], Thorsten Staudt [3] and Frank Mücklich [1,2]

1   Department of Materials Science, Saarland University, 66123 Saarbrücken, Germany
2   Materials Engineering Saarland (MECS), 66123 Saarbrücken, Germany
3   Aktien-Gesellschaft der Dillinger Hüttenwerke, 66763 Dillingen, Germany
*   Correspondence: bjoernivo.bachmann@uni-saarland.de

Abstract: Reliable microstructure characterization is essential for establishing process–microstructure–property links and effective quality control. Traditional manual microstructure analysis often struggles with objectivity, reproducibility, and scalability, particularly in complex materials. Machine learning methods offer a promising alternative but are hindered by the challenge of assigning an accurate and consistent ground truth, especially for complex microstructures. This paper introduces a methodology that uses correlative microscopy—combining light optical microscopy, scanning electron microscopy, and electron backscatter diffraction (EBSD)—to create objective, reproducible pixel-by-pixel annotations for ML training. In a semi-automated manner, EBSD-based annotations are employed to generate an objective ground truth mask for training a semantic segmentation model for quantifying simple light optical micrographs. The training masks are directly derived from raw EBSD data using modern deep learning methods. By using EBSD-based annotations, which incorporate crystallographic and misorientation data, the correctness and objectivity of the training mask creation can be assured. The final approach is capable of reproducibly and objectively differentiating bainite and martensite in optical micrographs of complex quenched steels. Through the reduction in the microstructural evaluation to light optical micrographs as the simplest and most widely used method, this way of quantifying microstructures is characterized by high efficiency as well as good scalability.

Keywords: quantification; microstructure; steels; correlative microscopy; machine learning; semantic segmentation

1. Introduction

In modern materials science, there has been a paradigm shift away from pure process–property correlations. Here, the microstructure has gained importance as a central key point for understanding the relationships between manufacturing processes and the resulting material properties [1]. The final microstructure can be seen as the signature of all influencing factors, including chemical composition and mechanical and thermal treatment [2]. In order to understand these different influencing factors, an objective and meaningful analysis and, above all, quantification of the microstructural components are essential. Furthermore, a reliable process and quality control depend on a well-founded evaluation of the microstructure [3]. This makes a reliable assessment of the microstructure an everyday problem in many quality laboratories and for process optimization in research and development. Particularly in these applications, the methodology for quantifying microstructures is severely limited in terms of time and cost. For this reason, classic light microscopic evaluations are most commonly used [4,5]. A light-optical microscopic (LOM) examination can be integrated quickly, easily, and cost-effectively into the quality control process [6].

However, the LOM method has limitations due to its limited resolution. As materials become more sophisticated and their microstructures become smaller and more intricate, comprising a combination of different phases or constituents with distinct substructures, tolerances narrow and quality requirements intensify. This has led to the limitations of existing characterization methods, especially in the case of LOM. Furthermore, sample preparation and especially sample contrasting using appropriate etching methods can have a major influence on the appearance of the corresponding micrographs [7]. Commonly used methods for the evaluation of LOM micrographs are based on manual estimations by experts. In view of the high complexity of the task, these expert opinions are thus characterized by a strong influence of subjectivity and are, therefore, directly dependent on the training and experience of the respective metallographers. Mostly, the expert's assessment can even just be reduced to a qualitative statement rather than a quantitative evaluation. Accordingly, in the context of the hereby presented showcase of quenched steels, it is usually the case that the expert assessment is limited to a statement of a "bainitic–martensitic microstructure" without any respective phase fractions.

Even partially automated methods of traditional image processing [8,9], including image segmentation based on different threshold values, are quickly limited in their effectiveness when dealing with images containing microstructural components that do not show a sufficiently distinct contrast. This is particularly the case when sample contrast and image acquisition conditions are not optimal or cannot be reproduced consistently or when it comes to complex microstructural types.

In addition to LOM, however, there are other promising methods that help to better quantify the microstructural constituents. It is worth noting the particular value of high-resolution scanning electron microscopy (SEM) [10] and electron backscatter diffraction (EBSD) in this context [11,12]. The high resolution of the SEM allows for the revelation of more details, which, in the case of steel, enables a more reliable classification in the usual classification schemes [13]. These schemes are typically based on fine morphological differences and the corresponding existence and/or orientation of the carbides present. Depending on the size and extent of these characteristics, the corresponding allocations can only be assumed based on the observation in the LOM. It is evident that classification problems may arise in which even high-resolution SEM images are unable to provide an unambiguous classification. This may be attributed to a number of factors, including orientation-dependent etching characteristics [14] and the fact that some classification schemes contain classes in which the transitions to other classes are often blurred and continuous in practical application due to generalized and simplified criteria [15]. For this reason, some classification schemes have been supplemented with crystallographic EBSD information [16] with the aim of creating a clearer boundary through complementary misorientation information. The formation of the respective microstructural components often leaves its mark in the form of characteristic misorientations, which contribute to the resulting mechanical properties [17]. Here, in particular, correlative microscopy represents a promising solution.

The principle of correlative microscopy is the combination of various scale-bridging experimental methods, with the aim of supplementing any missing information and thus simply benefiting from the strengths of the different investigation methods while excluding the disadvantages of others [18]. There is a large number of studies that benefit from this approach in order to enable a holistic and, above all, objective characterization of highly complex microstructures. The approach of using a combination of crystallographic EBSD data, high-resolution SEM images, and corresponding application-oriented LOM images to obtain the necessary information for unambiguous identification of the individual microstructural constituents has already proved to be successful at this point. The goal is to make an objective and correct classification based on correlative data and, where possible, reduce this to the most practical and efficient methodology, ideally involving the use of LOM micrographs.

Meanwhile, the use of machine learning-based (ML) computer vision (CV) approaches to characterize and quantify microstructural images is becoming increasingly popular [19–26]. This is mainly because ML-based approaches are able to deal with a wide range of variations within the micrographs after appropriate model training [27]. ML-based segmentation methodologies thus offer an auspicious alternative to traditional image processing techniques. Moreover, their high level of robustness and reproducibility obviate the issues associated with conventional manual evaluation by an expert, thereby creating a compelling alternative [28].

A well-known problem in the application of CV approaches to complex microstructural images is the determination of the ground truth [29]. Due to the high complexity of the image caused by a large number of influencing factors and the resolution limit of the LOM, it is often not possible to determine the ground truth unambiguously using only the LOM images on which the (serial) analysis is to be performed. At this point, the disagreement regarding the classification of graphite morphologies in cast iron, for example, is mentioned as a precedent, whereas the disagreement within the steel community increases accordingly due to the even higher complexity.

This is where correlative microscopy can unfold its strengths. The approach here is to generate a unique, high-quality set of correlative data from which an unambiguous ground truth can be determined. This ground truth is then used in combination with the simplest possible methodology to train an ML model that can be used in practical applications. Müller et al. [20] provide an impressive overview of application examples where correlative microscopy is used in combination with ML approaches to solve highly complex microstructure evaluation problems. In this way, the substructure of second-phase objects in dual-phase steel was quantified with the aid of correlative SEM images. With the information obtained from this, a classifier could be trained using traditional ML methods, which were able to distinguish pearlite, martensite, and bainitic second-phase objects from each other using morphological, textural, and substructural information.

Furthermore, Durmaz et al. [22] were able to establish a deep learning (DL)-based approach that could distinguish lath-shaped bainite from polygonal ferrite using LOM micrographs of complex phase steels. In this case, correlative SEM as well as EBSD data were consulted to clearly separate the two microstructural constituents based on the characteristic misorientations during the annotation process.

Bachmann et al. [30] and Laub et al. [31] used correlative LOM, SEM, and EBSD data to enable unambiguous segmentation of prior austenite grain boundaries in quenched steels. Using a combination of crystallographic EBSD reconstruction and manual corrections based on the visible austenite grain boundaries in the high-resolution SEM images, they were able to produce an objective and meaningful ground truth. These ground truth masks were used to train a model for semantic segmentation of optical micrographs, which gave impressive results based on picric acid-based etchings [31] and even based on common Nital etchings [30].

Bachmann et al. [32] extended the reproducible quantification of microstructural constituents to quenched and tempered steels, which have a much higher complexity than the other steel grades. The problem with these microstructures is that many image areas, despite the use of correlative SEM and EBSD data, do not allow a clear and unambiguous assignment to one of the common microstructural constituents based on their visual appearance. The authors selected a set of reference patches of the occurring phases to quantify entire micrographs using a sliding window classification technique in combination with a confidence threshold in order to identify microstructural constituents and to reproducibly quantify ambiguous areas in LOM as well as in SEM images.

Breumier et al. [33] trained a U-Net model for semantic segmentation that directly uses image-like numerical EBSD data (EBSD image quality, characteristic misorientations, and kernel average misorientation), as well as metadata to distinguish ferrite, martensite, and bainite phases in multiphase steels. In contrast to the quenched and tempered steels, a sufficiently reliable determination of the ground truth in the form of pixel-by-pixel annotation was possible for the majority of the microstructures. A distinguishing feature

of their methodology is that these specific areas were withheld throughout the training process, where no sufficiently reliable assignment could be made. These areas, therefore, have no negative influence on the model training and are assigned to the most probable class in the subsequent model application (=inference).

Furthermore, Shen et al. [34] succeeded in creating an EBSD-trained DL model that is able to accurately classify and quantify complex microstructures based on SEM images. The EBSD approach enabled the objective creation of a ground truth that was used to derive correlations between SEM and respective EBSD data.

The following work builds on the research mentioned above. The aim of this approach was to develop a model for the semantic segmentation of quenched and tempered steels that is capable of reliably segmenting, pixel-by-pixel, highly complex microstructures in a reproducible manner using only Nital-etched LOM images.

For this purpose, complementary information from LOM, SEM, and EBSD is used to clearly and objectively identify martensitic and bainitic areas. Based on these reference areas, a first model is trained that is capable of fully segmenting EBSD data. The result of this EBSD segmentation is then used as ground truth to obtain another model for the segmentation of martensite and bainite based on the correlative LOM micrographs. Through this automated mask generation, an objective ground truth can be created, which is based on objective crystallographic data and thus independent of etching and subjective assignments based on morphological characteristics through the annotating expert. Accordingly, the final model is able to reliably segment LOM micrographs without the need to collect correlative data for series application.

2. Materials and Methods

2.1. Material

The sample material is low-carbon steel (0.22% C) that was heat-treated using a quenching dilatometer (austenitized at 1000 °C for 10 min), which is part of the dataset used in [32]. For this work, 8 of the samples exhibiting various martensitic–bainitic microstructures (continuous cooling rates of 8–20 K/s), as well as respective samples representing purely martensitic (continuous cooling rate of 278 K/s) and purely bainitic (cooling at 50 K/s to 525 °C and isothermal holding for 300 s), microstructures were selected.

2.2. Correlative Microscopy

As mentioned above, it is difficult to identify all the microstructural components with high confidence from LOM images alone due to the high complexity and the distinct fineness of the microstructures investigated. Hence, correlative microscopy is used on all selected specimens to increase confidence in the microstructural classification. Furthermore, correlative EBSD and respective LOM data are necessary to apply the methodology presented here, as objective masks are first created by EBSD segmentation and then used as ground truth for segmentation based on LOM images.

For sample preparation, all specimens were mounted conductively. Subsequently, they were ground using silicon carbide paper with a grit size ranging from 80 to 1200 and polished with diamond suspensions of varying sizes, namely 6, 3, and 1 μm. A representative ROI was then marked using hardness indentations, which were used to perform the correlative microscopy routine.

EBSD was selected as the first method, as it requires a polished and non-contrasted sample surface. If one were to start with LOM or SEM images instead and carry out the EBSD measurement afterward, the polishing plane would be shifted by a corresponding back polishing to remove the contrasted layer, which would lead to considerable problems during subsequent image registration. Image registration refers to the targeted distortion of the individual microstructural images in order to make them congruent [34]. Due to the different contrast creation mechanisms of LOM, SEM, and EBSD, the corresponding identical areas cannot simply be superimposed in order to obtain congruent datasets [18].

As a preliminary preparation step for the EBSD measurement, the samples were prepared immediately before the measurement using colloidal OP-S polishing in order to completely remove fine, remaining scratches and deformation layers. Appropriate care and reproducibility during sample preparation are decisive for the quality of the EBSD measurement and, accordingly, for all subsequent results. The direct correlation between insufficient sample preparation and EBSD measurement quality has been demonstrated in a number of studies [35,36]. Furthermore, measurement settings of the EBSD and corresponding optimization of the focus of the electron beam and the stigmator, as well as calibration, background collection, and indexing settings, are crucial [37,38]. As the further workflow and the generation of the ground truth depend on the EBSD data, this process step can be identified as severely critical and decisive and is therefore kept constant after initial optimization. The $400 \times 400\ \mu m^2$ EBSD measurements with a step size of 0.35 μm were performed on a Zeiss Merlin (Zeiss, Oberkochen, Germany) with a corresponding EDAX EBSD detector at an accelerating voltage of 25 kV and a beam current of 10 nA at a working distance of 15 mm.

The EBSD data were post-processed using EDAX OIM 8 (EDAX, Pleasanton, CA, USA) to export the corresponding EBSD maps (Image Quality (IQ), Confidence Index (CI), Kernel Average Misorientation (KAM), and other misorientation information in the form of characteristic boundaries ($>2°$, $>5°$, $>15°$)). Therefore, a CI cleaning of values smaller than 0.01 was performed as a well-established common practice within OIM. These EBSD mappings were primarily used to register the LOM and SEM recordings and for complementary information during the annotation process but were not used for later application in the EBSD U-Net model for segmentation. For this purpose, the numerical EBSD data, which are available as measurement data in a list form, were converted into a matrix using the MATLAB (MathWorks, Natick, MA, USA) Toolbox MTEX 5.8 [39] and saved as a .mat file. Since the EBSD data were recorded in a hexagonal grid, it is necessary to partially interpolate them in order to convert a matrix to the cartesian coordinate systems corresponding to those of image data.

After successful EBSD measurement, the samples were polished again with OP-S to remove the contamination layer of the electron beam as well as oxidation layers in order to obtain a sufficiently good etching result using a 2.5% alcoholic Nital solution for 25 s. Corresponding LOM and SEM images of the region of interest were then captured. LOM images were taken at 1000× magnification (pixel size 126.6 nm) and stitched together to map the entire ROI using an Olympus LEXT OLS 4100 (Olympus, Shinjuku, Japan) for LOM. In the case of SEM, the micrographs were acquired using a Zeiss Supra (Zeiss, Oberkochen, Germany) SEM using the secondary electron contrast at an acceleration voltage of 5 kV and a magnification of 850× with a respective pixel size of 47.5 nm. Again, the individual images were stitched together using Microsoft Image (Version 10.2021) Composite Editor (Microsoft, Redmont, WA, USA).

LOM and SEM images were registered to the EBSD data, as these are the most complex and, therefore, the most difficult to manipulate, especially considering that the EBSD data are used directly in an ML workflow. The IQ map was used for image registration because it is the closest to a contrasted microstructure image by showing the corresponding topographical differences [18]. Image registration was carried out in Fiji [40] with the bUnwarpJ plugin [41] using manual feature selection. For this purpose, a selection of correspondingly identical sample points was chosen, on the basis of which the elastic transformation matrix for image registration was calculated.

The complexity of the underlying task of distinguishing martensitic from bainitic areas in this type of structure is particularly evident from the LOM images (Figure 1c). Often, there is no clear distinction between bainitic and martensitic areas, and particularly fine bainitic areas make clear characterization difficult. Due to the strong interconnection and interweaving of the two structural components based on the respective formation mechanisms, it is also hardly possible to define clear boundaries between bainitic islands in the predominantly martensitic microstructure. This is aggravated by the orientation

dependence of the Nital etching, which means that martensitic areas with a corresponding <100> orientation often appear bright and thus can be confused with other phases, such as bainite or retained austenite [32]. At those points where phase identification by contrast or the resolution limit of the corresponding method is limited, the complementary crystallographic (misorientation) information may provide support to reliably identify corresponding microstructural areas.
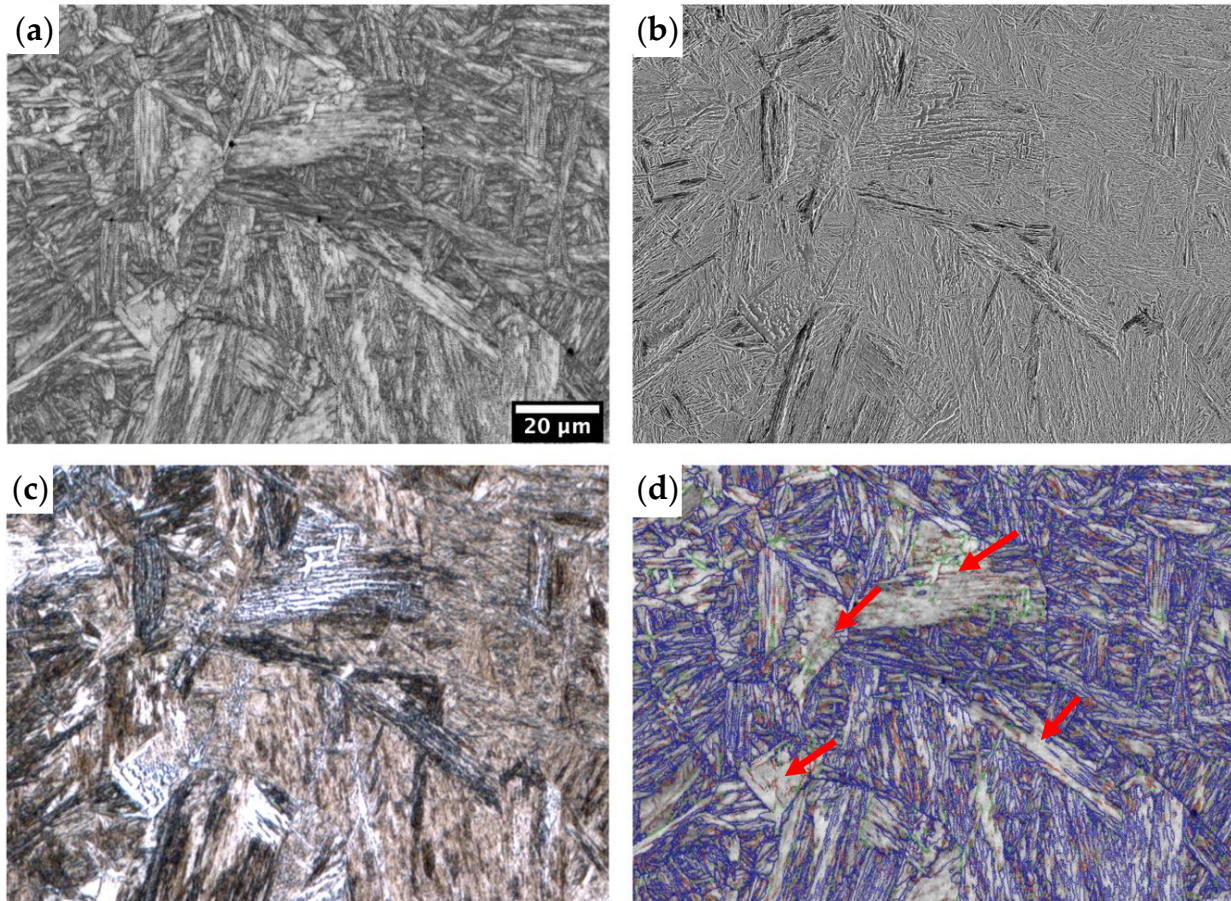

Figure 1. Representation of the same section of a bainitic–martensitic microstructure using correlative microscopy. Shown are the EBSD IQ (**a**), the corresponding SEM (**b**), as well as LOM (**c**) image and misorientation information (**d**) using characteristic boundaries (>2°—red, >5°—green, >15°—blue). Bainitic areas can be better distinguished from their martensitic counterparts, particularly due to the complementary EBSD information. Bainitic domains tendentially have a higher IQ (=brighter) and lower misorientation densities (red arrows).

2.3. EBSD Data as Objective Tools for Phase Identification in Complex Steels

The IQ metric describes the quality of an EBSD diffraction pattern by quantifying its contrast [42]. The perfection of the crystal lattice has a strong influence on the resulting imaging quality. Elastic lattice distortions, such as those found in martensite, displace the characteristic lattice planes [37]. As a result, the exact Bragg condition is no longer met. This results in blurring of the Kikuchi bands, which leads to lower IQ values. The same applies to the EBSD pattern of grain boundaries or precipitated phases, such as cementite. Accordingly, bainitic ferrite with moderate misorientation tends to have higher IQ grades, and dislocation-free polygonal ferrite achieves the highest IQ grades.

The situation is similar to the CI. The CI does not quantify the quality of the pattern itself but the confidence with which the measured pattern could be indexed by comparing each pattern individually with a crystallographic database [43]. Analogously to IQ, increasing dislocation density and precipitated second phases, such as carbides and MAs, cause a

poorer CI when comparing the indexing to an iron-focused base-centered cubic database in the case of steels and thus can act as a distinguishing criterion between, e.g., polygonal ferrite with low dislocation densities and characteristic bainitic morphologies, as well as martensitic microstructural components [15,44].

The KAM value quantifies misorientations by means of local orientation gradients. For this purpose, the mean misorientation to the central measuring point is calculated within a kernel of a certain order. In contrast to the grain average misorientation (GAM) and the grain orientation spread (GOS), which are both object-based values and therefore require an object/grain definition [45], the KAM is a pixel-based value. This makes the KAM easier to generalize and more meaningful with regard to local misorientations. This is especially the case when it comes to martensitic microstructures, where no analogous grain definition can be conducted due to the lath/needle-like character. The characteristics of the KAM values can, therefore, also be used to distinguish occurring microstructural constituents in steels, like low-dislocation polygonal ferrite, from different types of bainite with, respectively, unique manifestations of the KAM, as well as martensitic constituents with exceptionally high KAM values [15].

Due to the sufficient information density as well as the good comprehensibility and interpretability of the underlying pixel-based data of IQ, CI, and KAM, the scope of this work has been reduced to these three parameters.

There are several works that differentiate between different structural components in steels based on various EBSD parameters using various thresholding or multi-conditions [46–51]. A common feature of all these works is that they use a similar selection of EBSD parameters, particularly mentioned IQ, CI, and KAM. These parameters are then traced back to the corresponding structural development and the resulting local misorientations of the respective microstructural constituents in order to identify and, afterward, quantify them. Most of the work is primarily devoted to the quantification of occurring microstructural constituents of dual-phase and multiphase steels. At this point, however, the authors are not aware of any work that deals with the quantification of highly complex quenched steels based on an approach comparable to the aforementioned. Accordingly, this work builds on the basic principles introduced to differentiate between martensite and bainitic constituents in highly complex quenched steels.

The LOM and even the high-resolution SEM images clearly show the complexity of the investigated microstructures and the underlying challenge that arises, attempting to fully characterize these microstructures using only LOM images. Although some clear areas can be identified (colored arrows in Figure 2), there are also many areas that cannot clearly be classified. The characterization of the transition areas, in particular, poses a major challenge. However, the criteria developed to differentiate between martensitic and bainitic areas can be identified relatively clearly. Thus, the bainitic areas have higher IQ and CI values (bright) with simultaneously lower misorientation, expressed by the KAM map, than their martensitic counterparts. Based on the EBSD data, the majority of this microstructure section can, therefore, be assigned to one of the considered classes (regions at the colored ellipsis in Figure 2). Specifically, when it comes to determining the boundaries of the respective structural components, consulting the EBSD maps helps enormously. However, the sum of all correlative data from EBSD, LOM, and SEM does not allow a clear assignment of all pixels/measurement points. In some areas, there is a certain amount of room for interpretation due to the partly continuous transition between the two microstructures caused by the continuous cooling, which was intentionally tried to avoid. The challenge now is to deal with this uncertainty in a proactive way while taking advantage of the benefits of modern ML approaches, as outlined above.
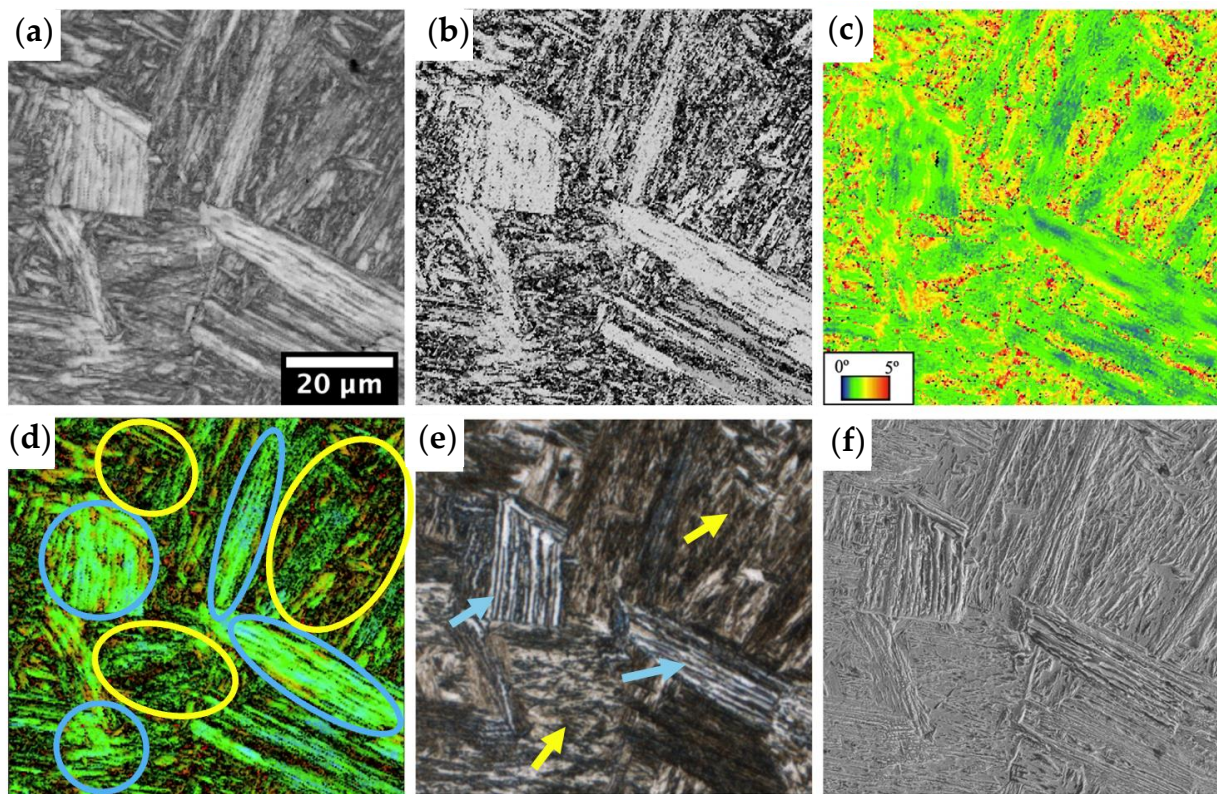
Figure 2. Correlative recordings of a highly complex quenched steel with bainitic (blue) and martensitic (yellow) regions (in (**d**,**e**)). Correspondingly isolated EBSD measured variables IQ (**a**), CI (**b**), and KAM first order (**c**), as well as the corresponding superposition of these three (**d**) in addition to the corresponding LOM (**e**), as well as the SEM (**f**) image after corresponding contrasting of the microstructures with Nital.

2.4. Semantic Segmentation of EBSD Data

As previously stated, a substantial corpus of research exists on the quantification of various microstructures based on EBSD data. However, this is predominantly limited to dual-phase or multiphase steels. In such cases, it is possible to differentiate between them using simple threshold values or multi-condition techniques. Nevertheless, these approaches encounter limitations when applied to more complex microstructures. For instance, a simple threshold value segmentation is not sufficient to fully segment bainitic areas. In bainitic areas in particular, both high and low IQ/CI values frequently occur, and there is often no homogeneous distribution of these values due to the existence of characteristic carbide or (sub-)grain boundaries [15]. The situation is similar with martensitic microstructures. The martensitic microstructure is particularly characterized by poor indexing, thus low IQ and CI values, high KAM values, and a needle-shaped morphology. However, depending on the fineness of the microstructure and the related step size of the EBSD, relatively well-indexed martensite laths can also exist. Simultaneously, a strong orientation influence can be observed [32]. Furthermore, the orientation of the martensitic packages to the grinding plane can lead to a variety of different morphologies.

For these reasons, the use of semantic segmentation is particularly suitable for the problems addressed. Semantic segmentation is a process whereby an entire input image is classified pixel by pixel. The advantage of this approach is that it enables the deep learning (DL) model to understand semantic information and place it in context with surrounding areas [52]. Consequently, the image understanding of semantic segmentation approaches is sufficient to solve highly complex segmentation problems where conventional approaches fail. Therefore, the distinguishability of different classes is not limited to conventional features, such as homogeneity of color values or similarities.

Semantic segmentation can be applied to all image-like data, thus also to EBSD data. While Ostormujov et al. [53] and Breumier et al. [33] limit the application of semantic segmentation to EBSD data of two-phase or multiphase steels, where the annotation leaves only little room for error, the following approach is aimed at the quenched steels presented. As already described, the problem with these is that not all areas allow clear phase assignments. However, this is exactly the point where the approach presented here comes into play. The aim is to use a set of reference areas to train a U-Net for segmenting EBSD data that can objectively and reproducibly separate martensitic from bainitic areas. Thus, the allocation is no longer based on the qualitative assessment by the experts by setting a corresponding threshold or approximate estimation but on the reproducible application of the model to the entire EBSD data. Another advantage of using semantic segmentation is its probabilistic character. The result of semantic segmentation contains the absolute probability of each individual class learned. Accordingly, areas where there is residual uncertainty can be declared by a corresponding confidence threshold and thus taken into account.

Due to the residual uncertainty shown despite correlative data with regard to some areas, the data are only partially annotated. Only pixels, where morphological features and the appearance of each individual method, correlative LOM, SEM, and qualitative EBSD data agree towards a microstructural class, martensite or bainite, mask annotations were created. Accordingly, areas for which neither a clear martensitic nor bainitic affiliation could be identified are summarized in a common "background class." Hence, the annotations can be objectively assumed to be consistent as well as correct in order to use them in a cross-methodology DL approach. This need for only partial ground truth masks also speeds up the annotation process enormously. A similar approach, but with the aim of creating a more efficient annotation routine, was presented by Na et al. [54]. After appropriate modification of the loss function, the unlabeled areas are neither positively nor negatively included in the training process. The loss function is a numerical representation of the objective that must be minimized during training. A suitable selection of the loss function is crucial to achieve the respective segmentation objective [55]. Nevertheless, such areas can be segmented in the inference by assigning them to the class that the model considers most likely according to the objective learned based on the training data.

Only those areas were annotated for which the characteristics of the respective class affiliation within the individual methods of correlative microscopy match. Thus, the influence of etching and measurement artifacts, as well as a certain orientation influence, can be taken into account. An example of these critical points can be found in Figure 3c. For example, there is an area with less local misorientation (top right), which is indicated by lower KAM values. Based purely on the misorientation, the appearance of this area is more similar to that of bainite (see blue areas). However, based on IQ, CI, LOM, and SEM, this area can be clearly identified as martensitic. At this point, the great advantage of correlative microscopy once again becomes clear, as it can significantly increase objectivity and identify corresponding anomalies in the complex structures that would otherwise have led to incorrect annotation.

The EBSD training dataset consists of 8 measurements with the dimensions 1320×1320 measuring points. Each measurement field is divided into individual patches of size 256×256 px before training, resulting in 200 total patches and corresponding masks. A train test split of 20–80 was then applied.
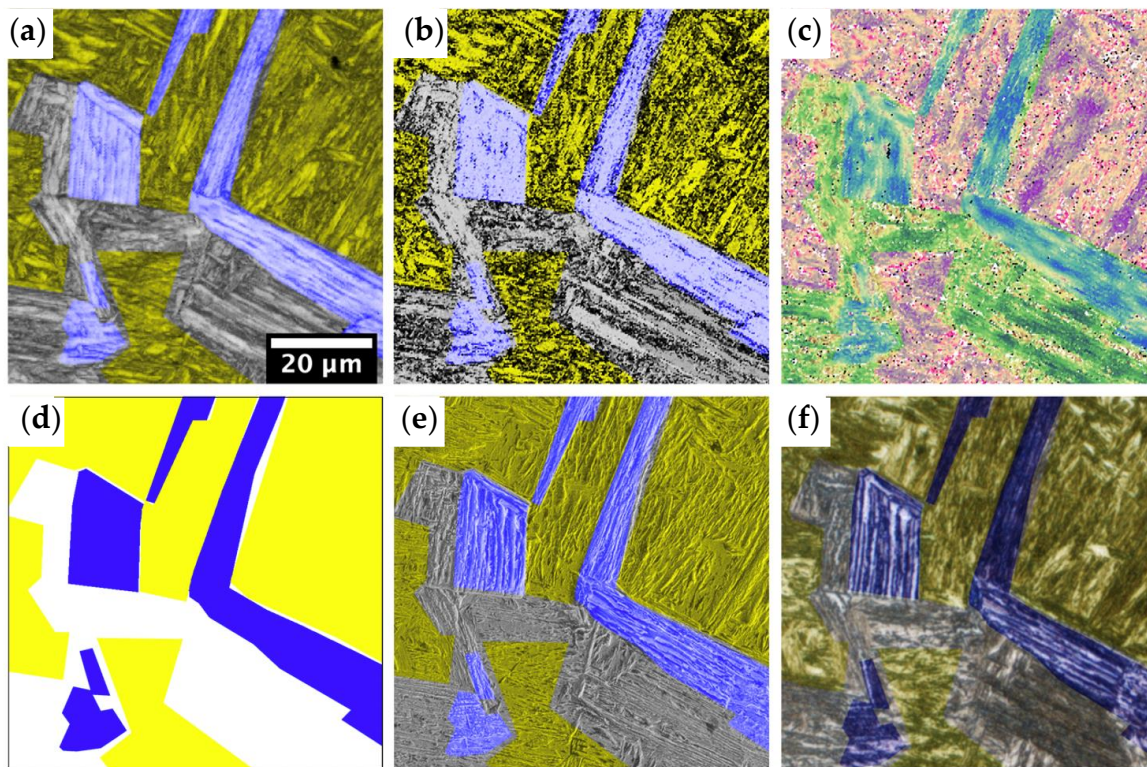
Figure 3. The partial annotation (martensite—yellow, bainite—blue, unlabeled—white) (**d**) as overlays with respective EBSD mapping ((**a**) IQ, (**b**) CI, (**c**) KAM (colors of the mask adjusted for clarity—red tone as result of the overlay of the yellow mask with the KAM colorcode)), as well as with micrographs (SEM (**e**) and LOM (**f**)). Shown is the same sample section as in Figure 2.

The proven U-Net—from the segmentation models package [56]—was chosen as the model architecture. The Inceptionv3 [57] with ImageNet pre-trained and trainable weights was chosen as the encoder part. IQ, CI, and KAM (first order) data were selected as input channels. Since absolute IQ data can only be compared with each other to a limited extent due to the high sensitivity of the EBSD measurement methodology to calibration [37,38], each EBSD dataset was standardized individually in this respect. CI values are already available in a normalized form (0–1). The KAM values were capped at a local misorientation of $5°$ and normalized correspondingly. In total, the entire dataset contains 12% bainitic, 23% martensitic, and 65% unlabeled areas. Class weighting (only between the bainitic and the martensitic classes) was integrated in order to reduce objective and possible bias caused by the class imbalance. A modified categorical cross-entropy was used as the loss function, which only focuses on the annotated areas (bainite and martensite). The final prediction step incorporates a softmax function in order to integrate the probabilistic character into the final result. Due to the comparatively high proportion of unlabeled data that is not considered in the loss function, as well as the data imbalance between bainitic and martensitic ranges, a class-based IoU that is as comparable as possible between the two classes is desirable. This way, a bias in favor of one of the two classes can be reduced, and the corresponding unlabeled data can be segmented as objectively as possible. The model was trained for 10 epochs using an Adam Optimizer at a learning rate of 0.0001. The model was then fine-tuned with a correspondingly small step size until the metric intersection over union (IoU) of the two classes were as close to each other as possible and converged. To increase the amount of data, the data were slightly augmented ($90°$ rotation and flipping). No further augmentation, e.g., contrast or brightness, was applied to avoid unintentional falsification of the EBSD data (Figure 4).
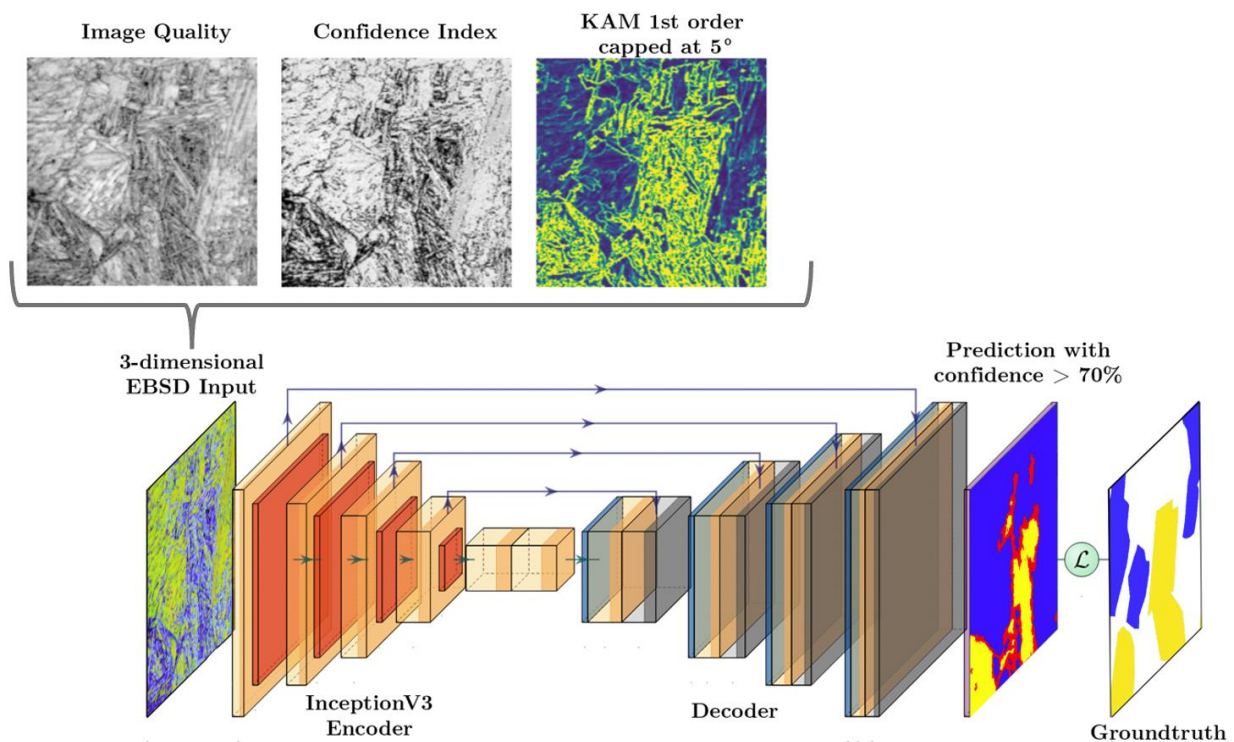
Figure 4. Schematic illustration of the used U-Net architecture [modified [58]] with ImageNet pre-trained InceptionV3 encoder. Each individual EBSD channel (IQ, CI, and KAM) is concatenated to an RGB-like image after normalization to feed into U-Net. During the training process, the model learns the characteristic features of each individual class (martensite—yellow and bainite—blue). The white, unlabeled regions from the ground truth masks do not influence the learning process. For inference, the segmentation results are filtered by a confidence threshold of 70% (red) in order to create more meaningful results. As visible, the model is able to predict unlabeled regions. However, the loss function "L" only considers the labeled pixels during calculation by comparing the unfiltered prediction and ground truth mask during training.

2.5. LOM Segmentation Pipeline

As already motivated, a well-engineered segmentation pipeline based on EBSD data can provide an objective and reliable segmentation of microstructural images. Due to the functional principle of EBSD measurement, these assessments are thus based on valid and objective crystallographic information. Based on this, microstructural constituents can be distinguished from one another. The resulting objectivity is utilized, and the segmentation results are used as a starting point for a further approach to segment LOM images. This not only means that ambiguous areas can be labeled objectively but also that training masks can be created extremely efficiently for highly complex microstructures where manual labeling is very time-consuming. Considering the previously mentioned amount of unlabeled data (about 65%), far more training data can be generated by applying the learned criteria to the entire dataset.

As briefly mentioned, only the highly confident predictions are used as labels to train the LOM segmentation pipeline. Thus, a respective measurement/image point must belong to the respective class with at least 70% probability in order to be used as a ground truth. All filtered predictions for which the EBSD segmentation model is not correspondingly confident are again counted in the "unlabeled" class. In addition, a median filter is used to smooth the predictions and to reduce noise or artifacts (Figure 5).
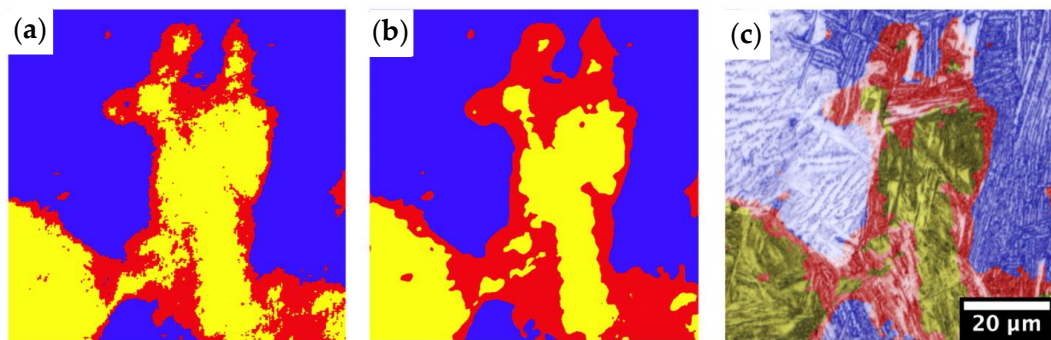
Figure 5. Segmentation result based on EBSD data after applying the 70% confidence threshold (**a**), with respective smoothing using the median filter (**b**). This image is subsequently used to train the segmentation model based on LOM images as an input, with blue representing bainite, yellow martensite, and red the unlabeled pixel, respectively. LOM is overlayed with an automatically generated training mask (**c**). No wrong labels could be identified using the correlative information. However, the confidence threshold of 75% is intentionally set conservatively in order to maintain a high level of objectivity and assurance in the labels. Red pixels correspond to confidence filtered regions.

Using the native resolution of the LOM images, the final dataset contained 7 correlative LOM images of size 4096 × 4096 px and 7 corresponding training masks coming from the post-processed prediction of the EBSD U-Net. One of the correlative datasets included in the EBSD segmentation approach was entirely excluded from the LOM training to use for further validation. The images were divided into 512 × 512 px patches, resulting in a total dataset of 448 patches. Once again, a train–test split of 20–80 was performed, and data augmentation just included rotations and flipping. The training images were preprocessed before training according to the standard procedure of the encoder with regard to ImageNet pretraining. For training the LOM segmentation model, the same architecture as for the EBSD approach was used. In this case, however, the weights of the encoder remained static. Since a superposition has proven to be beneficial as a loss function for segmenting complex LOM problems in previous works, [30] is used here as a guide. The training was performed for 10 epochs at a learning rate of 0.0001, with subsequent fine-tuning at a learning rate of 0.000001.

3. Results and Discussion

3.1. Semantic Segmentation Based on EBSD Data

The EBSD U-Net showed a class-wise IoU of 98.3% and 98.6% on the training set and 93.9% and 95.5% on the validation set, each for the bainitic and the martensitic classes, respectively. For semantic segmentation, these are extremely high metrics. From this, it can be concluded that the respective annotated microstructure sections, which are clear reference areas of the microstructure components, are recognized correctly for the major part. What is interesting at this point, however, is the extent to which the model generalizes to domains that do not have an annotation base.

At first glance, the segmentation model delivers excellent results, confirming the model metrics identified during training. The post-processing routine also does not falsify the results. Increasing the confidence threshold obviously leads to more uncertain areas. These uncertain areas are mainly located at the interfaces between the different phases. This was observed in [32] as well and can also be identified as the main reason why a complete annotation of the entire image area would be so hard to reproduce. In this way, this uncertainty or these transition zones can be reliably recognized and quantified. Using the 50% confidence threshold (Figure 6a) leads to 13% of pixels that cannot be assigned to a class. Increasing the confidence threshold to 70% (Figure 6d) increases this fraction to 29% total.
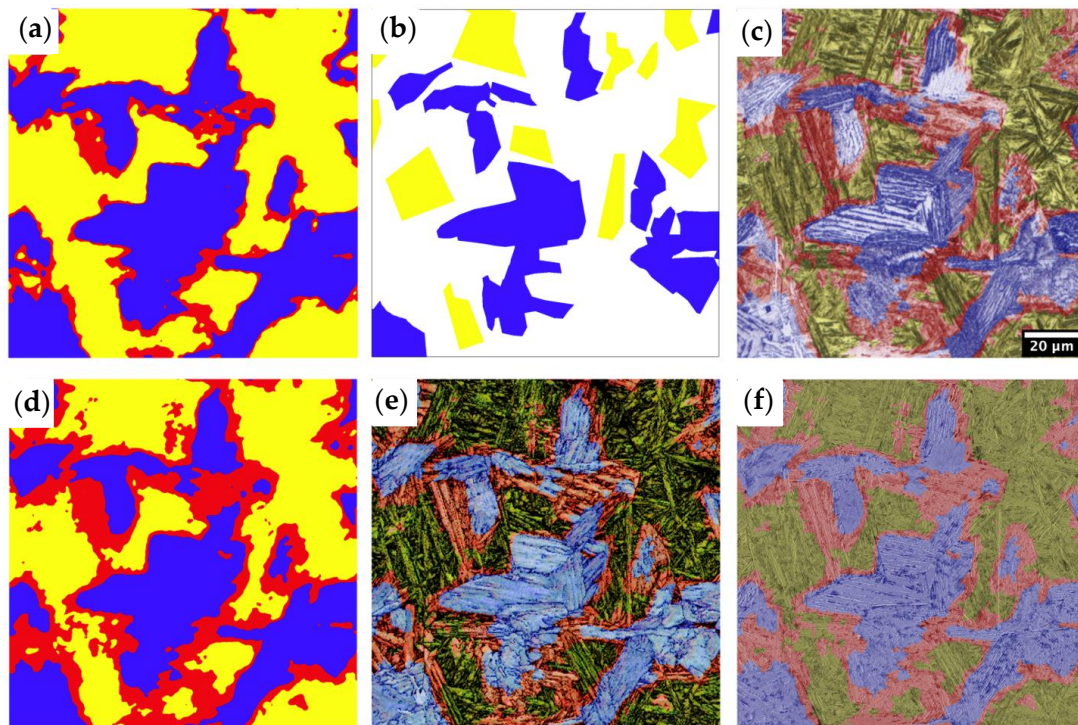
Figure 6. Segmentation results of the presented model using the EBSD data include a 0.5 confidence threshold (**a**) and 0.7 (**d**) after applying the median filter (bainite—blue, martensite—yellow, filtered—red). (**b**) shows the validation mask, which was partially annotated, and (**e**) as an overlay between the 0.7 threshold result and the EBSD input array consisting of IQ, CI, and KAM. (**c**) and (**f**) also show the post-processed segmentation result as an overlay with LOM and SEM, respectively.

However, there is little change in the relative proportion of the two phases. If only the reliably predicted proportions are considered for phase quantification, the model predicts phase proportions of just under 55% martensite and 45% bainite with stricter specifications. With the same calculation and a confidence threshold of 50%, the phase proportions are 57% martensite and 43% bainite. This finding suggests that the present model meets the expectations. The model is able to assign the majority of the microstructural areas to one of the two classes. By correcting the confidence threshold, the scope for interpretation can be set, which thereby quantifies the complexity of a structure and influences the number of domains that are more ambiguous than others. Based on this example, the threshold has only a negligible influence on the final quantification of the two phases. In particular, microstructures in which the transformation is not complete for the most part due to the continuous transition during continuous cooling could be quantified using this approach.

The bainitic areas (blue) stand out due to a few misorientations (light green to bluish) and high indexabilities (bright IQ and CI) (see Figure 7a). Martensitic areas, predicted in yellow (Figure 7b,c), show high misorientations (yellow to red), as well as some poorly indexed points (black) and a high (grain) boundary density (see Figure 7a). However, the assessment of the filtered areas marked in Figure 7a (red ellipsis) is less clear. These show comparably good pattern qualities or indexability at the EBSD level (bright IQ and CI) but possibly too pronounced local misorientations for bainite (yellow–red) (Figure 7a). When observing the morphology of the areas in question, ambivalent characteristics can also be identified. These areas consist of several lath-shaped packages. Depending on the type of bainite, this can fall under both the bainitic and the martensitic classes. These packets are comparatively narrow, which is more in favor of a martensitic microstructure component. However, precipitated carbides can be recognized within these lath packages, which appear in the EBSD in the form of parallel, poorly indexed "boundaries" but become clearer on the SEM (Figure 7c). There, the morphology of this area is reminiscent of extremely fine upper

bainite. At this point, the major influence of the selected step size of the EBSD acquisition should be mentioned. The finer the microstructures become, the poorer pattern quality or indexing occurs. Accordingly, the step size must always be adapted to the fineness of the underlying microstructure. However, since the SEM does not provide any clarification at this point either but merely reinforces the ambivalence of the areas examined, it can be concluded that this example represents the complex case of blurred boundaries and continuous transition between the two microstructural constituents. Accordingly, due to the continuous cooling, it is an ambiguous structure that cannot be clearly assigned to one of the two classes examined. Accordingly, it is easy to understand why the EBSD U-Net was unable to make a clear statement.
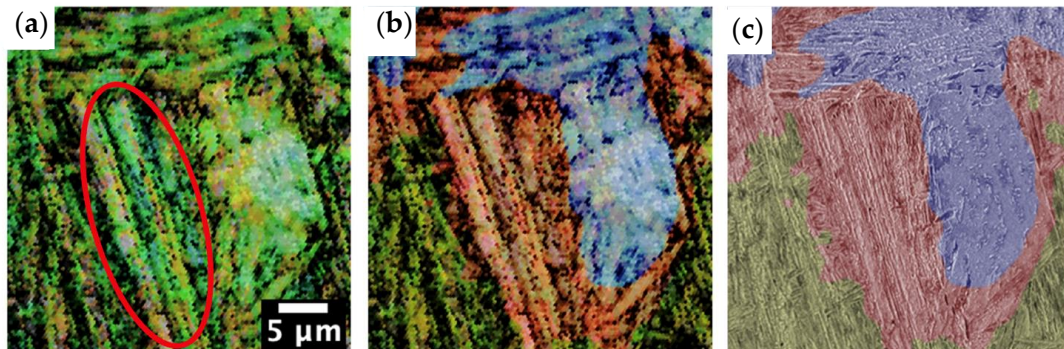

Figure 7. Magnification of an unclear area (red ellipsis) from Figure 6. Input array consisting of IQ, CI, and KAM (**a**), as well as overlay with post-processed prediction (**b**) (blue—bainite, yellow—martensite, filtered—red) and correlative high-resolution SEM image with corresponding overlaid prediction for clarification (**c**).

It can, therefore, be concluded that the conservative estimates with a confidence threshold of 70% of the EBSD segmentation model are suitable for training the LOM-based segmentation approach.

3.2. Semantic Segmentation Based on LOM Micrographs

The LOM-based U-Net achieved high segmentation metrics considering the underlying complexity and the limited resolution of the LOM. Based on the underlying training dataset from the seven samples, class-specific IoUs of 95% for the bainitic class and 94.4% for the case of the martensitic class could be achieved. In the case of the validation dataset, these values were 88.9% for bainite and 89.3% for the martensitic class. The balance of these values suggests a model that is not biased and, therefore, addresses both classes equally.

A direct comparison of the two segmentation results from EBSD (Figure 8d) and LOM (Figure 8e) shows a predominant agreement. Even fine bainitic areas could be successfully identified due to their coloration and similarity to the martensitic phase. However, the prediction of the LOM-based U-Net seems to have a higher confidence in general. For example, with the same confidence threshold (0.7), 30% of the pixels are labeled as uncertain for the EBSD prediction and only 15% for the LOM-based prediction in this shown microstructural domain. In view of the depth of information provided by the EBSD technique, this seems surprising at first but speaks in favor of the LOM-based evaluation. One possible reason for this could be the amount of training data. The EBSD-based model had to make do with fewer labeled data, whereas the LOM U-Net was provided with more training data due to the good performance of the EBSD model, creating the training masks. Another cause could be the color information stored in the LOM in addition to the topography. Due to different carbon contents, which are one of the characteristics of the different microstructure components, different color characteristics are created by contrasting using Nital. Through reproducible etching and a narrow parameter space, the DL model can understand correlations between color tone and microstructural constituents and apply them accordingly. In terms of resolution, the EBSD measurement with a given

step size of 0.35 μm is in the same order of magnitude as visible light. At this point, it would be interesting to see whether a finer step size would improve the performance and confidence of the EBSD-based model for phase analysis, which would be logical. However, due to the statistical nature of the learning process of DL approaches, it is not easily possible to compare the absolute confidence values of the two approaches.
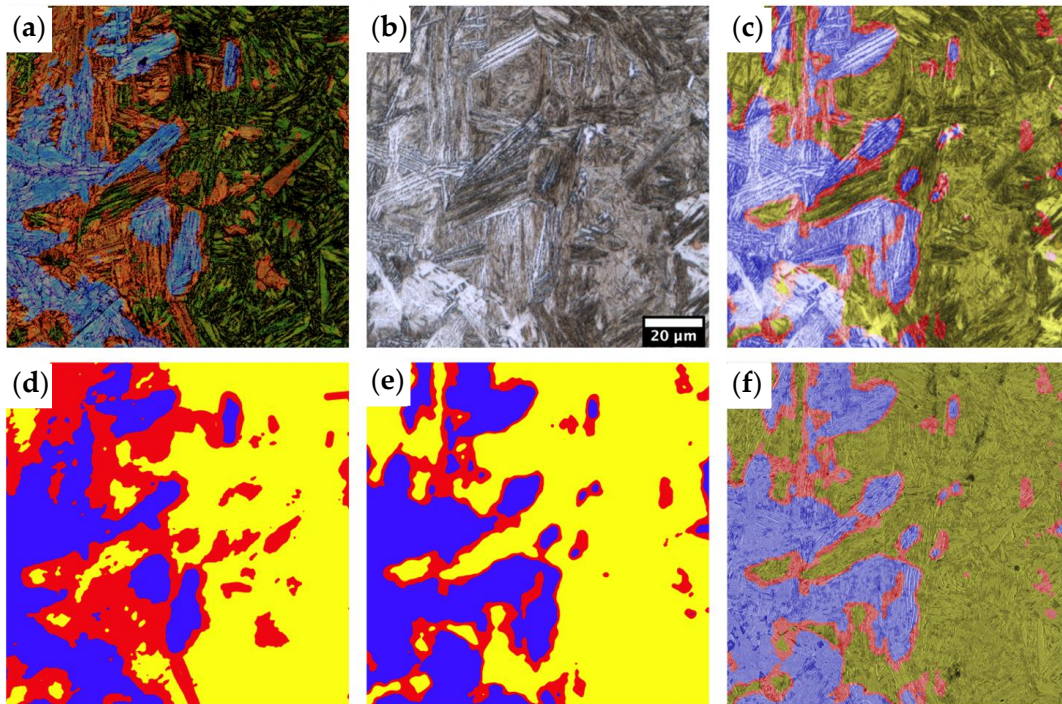

Figure 8. Correlative EBSD input array overlaid with the corresponding post-processed prediction of the EBSD UNet model (blue—bainite, yellow—martensite, filtered—red) (**a**). The corresponding mask (**d**) served as the ground truth for training the LOM images (**b**). (**e**) shows the post-processed (analogous to EBSD post-processing) prediction of the trained LOM UNet. (**c**) displays the overlaid representation of this mask with correlating LOM input for clarification. A better assessment of the LOM-based result can be made using the high-resolution SEM image (**f**).

However, a direct comparison of the predicted phase fractions between the EBSD and LOM variants can provide further information on the comparability of the two methods. If only the labeled pixels are considered, excluding the filtered areas (red), the respective phase fractions are very close to each other. Thus, the EBSD model determines a martensitic phase of 68% with a corresponding 32% bainitic phase. The martensite content determined using the LOM model is 68.6%, which corresponds to a bainite content of 31.4%. The misclassified pixels between the EBSD and LOM approaches, i.e., those that are assigned to the other class, excluding those that were filtered, represent 1.5% of the total pixels in this example. This also demonstrates the robustness and reliability of these approaches. Due to the remaining number of ambivalent sections, a clear ground truth mask of all pixels cannot be created manually in an objective and reproducible manner. Hence, quantifying the underlying error is not straightforward. Regarding the high-performance metrics and the advanced underlying methodology of obtaining the EBSD results, these are used as a benchmark in order to quantify a misclassification error.

However, these figures only apply to the section shown in Figure 8. The LOM evaluation of the sample (see Figures 2 and 3) that is included in the EBSD dataset but was not used to train or validate the LOM model is discussed in the following. This allows the robustness of the reduction in this metallographic task to the LOM as the simplest methodology to be further verified.

Again, the majority of the segmented areas of both approaches agree. Based on the LOM image (see Figure 9a), 90% of the section is segmented as martensite and 4.7% of the pixels as bainitic. No reliable statement can be made about 5.1% of the pixels. After correction by excluding the filtered pixels, this microstructure section consists of 95% martensitic and 5% bainitic components.
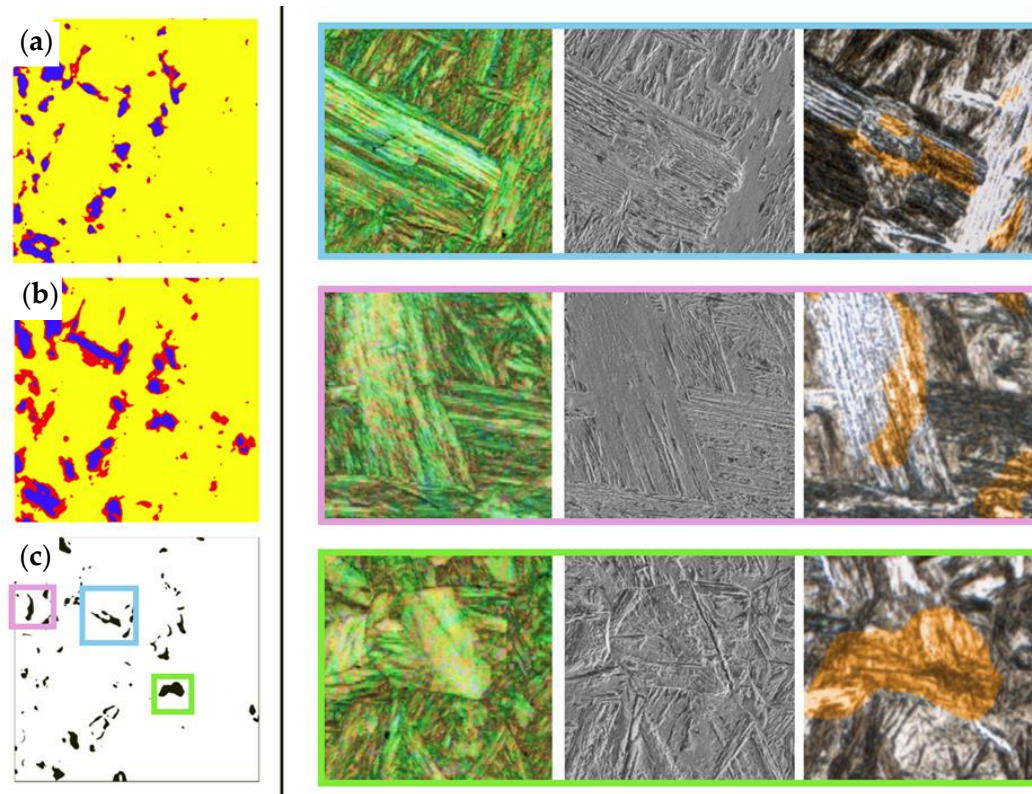

Figure 9. Shown are the segmentation of the LOM-based model (**a**) and the EBSD-based model (**b**) for the identical ROI of the withheld sample. (**c**) shows the oppositely classified pixels between (**a**) and (**b**) (orange). A closer look at the underlying correlative data (left EBSD, middle SEM, right LOM) of these misclassified areas (colored squares) provides information about possible reasons for the misclassifications. The discrepancy areas are marked in orange in the respective LOM.

The evaluation of the EBSD data shows a martensite proportion of 80.8%, a bainite proportion of 6.5%, and no certainty for 12.7% of the pixels. After normalization, excluding the uncertain portions, the microstructure, according to EBSD segmentation, is composed of 92.6% martensite and 7.4% bainite. Accordingly, the results of the two methods are again very close to each other. A total of 4% of the total pixels were misclassified (see Figure 9c), i.e., classified as bainite based on LOM and martensite based on EBSD, and vice versa. A direct comparison of the underlying EBSD, LOM, and complementary SEM data can be used to further investigate the oppositely classified areas.

In the misclassified area marked in blue (Figure 9), the majority of the bainitic package was not recognized on the basis of the LOM. The SEM image confirms the assumption that the segmentation based on the EBSD data can be assumed to be correct. Based on the LOM, however, the area in question (Figure 9, red ellipse) can be clearly identified as bainitic. Based on the dark coloration as well as the fine laths, the misclassification of the LOM at this point appears to be understandable. This case reflects the limitation due to the resolution limit of the LOM at this point, in combination with respective etching influences.

The situation is similar in the area marked in pink (Figure 9). However, this section allows a little more room for interpretation. Once again, the EBSD-based approach estimates this area more generously than the LOM evaluation. Neither LOM nor SEM or EBSD data

allow a clear allocation at this point. This means that there are definitely areas, albeit few, where an even stricter confidence threshold would be expedient.

The area marked in green (Figure 9) is a special case. Here, both LOM and correlative SEM images clearly indicate a martensitic microstructure. Nevertheless, the classification as bainite can be followed on the basis of the underlying EBSD data. The misclassified area shows a comparatively large and homogeneous area of high IQ and CI. The moderate misorientations (KAM) differ little from comparable actually bainitic areas. Nevertheless, these areas undoubtedly exhibit martensitic phenomena (LOM and SEM). This special case could, therefore, be an example of the aforementioned IQ–orientation relationship [37]. Assuming a favorably oriented martensitic package was cut parallel in this plane, this could result in the corresponding appearance of a homogeneous area of high indexing. Thus, despite its high objectivity, the EBSD cannot be said to be absolutely correct. Rather, it should be emphasized at this point that only correlative microscopy helps to characterize structures holistically and as objectively as possible. Each experimental method has its own strengths and weaknesses, and at the current state of the art, there is no method that is able to characterize highly complex structures with absolute certainty and at the same time with great efficiency and statistical significance.

4. Conclusions and Outlooks

In the present work, modern ML approaches were used to quantify the highly complex microstructures of quenched steels. A corresponding as objectively as possible, which represents a decisive hurdle in view of the complexity of the microstructures investigated, was established by means of correlative microscopy. For this purpose, LOM, SEM, and EBSD data were used in a complementary manner to identify clear reference areas of the two microstructural constituents, martensite and bainite. However, the problem with such microstructures is that not all areas can be annotated accordingly. Training a model for semantic segmentation is only possible by modifying the loss function, which means that non-annotated areas have no influence on the learning process during training. Nevertheless, in the subsequent segmentation, also known as inference, the model is able to assign these areas to the most probable class. A corresponding confidence post-processing routine was used to integrate uncertainties.

In the beginning, a U-Net model was trained that is able to perform phase discrimination based on the manually annotated reference areas using EBSD data (IQ, CI, and KAM were used for this purpose). Since the EBSD data have a crystallographic origin and can capture local misorientations qualitatively very efficiently, which have been proven to be a central decision criterion for phase identification, this approach can be regarded as objective and reproducible. For both classes, bainite and martensite, this model was able to achieve high IoUs of over 93% on the validation dataset. By integrating a correspondingly high confidence interval of 70%, the segmentation results could be given additional credibility.

These filtered masks were then used as ground truth to train another U-Net segmentation model, which was trained with correlative LOM micrographs. As with the previous training procedure, the filtered areas had no influence on the learning process of the model. This LOM model also achieved high class-specific IoUs of over 88%, respectively. Thus, the goal could be achieved by using correlative microscopy specifically to create an objective segmentation routine. However, this segmentation could be successfully reduced to the simplest possible experimental methodology, the LOM. Only this allows efficient evaluation and, thus, integration into quality control, for example.

Furthermore, it could be shown that the evaluations of the two evaluation routines, based on LOM or EBSD, differ only slightly with regard to the final phase proportions. Evaluation example 1 (Figure 8) resulted in a bainite content of 32%, with a corresponding martensite content of 68% based on the EBSD approach. The LOM approach yielded values of 68.6% martensite and 31.4% bainite. Example 2 (Figure 8), which includes an additional sample unseen by the LOM model, had an EBSD-based martensite content of 92.6% with a corresponding 7.4% bainite phase. The LOM model determined comparable values of 95%

and 5%, respectively. The number of pixels classified in opposite ways between the EBSD and LOM approaches was comparatively low at 1.5% (example 1, Figure 8) and 4% (example 2, Figure 9). These minor deviations from one another ultimately demonstrate the added value of the presented approach. In most practical applications, the microstructures dealt with here cannot be quantified at all, let alone reproducibly and objectively, with specific phase fractions. In this case study, however, the significance of the quantification based on the LOM recording could even be validated by the underlying objective EBSD data.

There are various approaches to color etching in the literature that are intended to make it possible to reliably distinguish between different structural components. Known problems that occur with color etching are, on the one hand, safety risks and, on the other hand, commonly poor reproducibility. Based on the approach presented here, reliable phase differentiation could be demonstrated using a simple and widely used Nital etching, which is very attractive for practical application.

The problem that preceded the work was solved by means of an appropriately developed methodology, and the reproducibility and objectivity of the approaches presented were demonstrated. The developed workflow thus provides the basis for the systematic creation of an objective evaluation routine for highly complex structures based on the simplest and most efficient methodology. This procedure can be extended accordingly to other microstructure types and material classes, as well as by integrating further microstructural (sub)classes. It can be assumed that this model is applicable to similar steels as used to create the training data, though it most likely is not strictly limited to this specific chemical composition. If previously unknown additional microstructural components or different forms of the known bainite class, for example, occur, it can be assumed that this model will have problems identifying them correctly. The reason for this lies in the necessary extrapolation of the ML model, for which these are usually not well suited. However, this arising problem can be tackled by extending the dataset by the occurring variations of the further microconstituents and retraining the model using the presented methodology to retain objectivity.

Thus, this approach is not only limited to steel microstructures but can be applied analogously to other material classes, e.g., complex titanium alloys. However, it should be weighed whether it is necessary to integrate time-consuming EBSD data or whether the use of more simple correlative information (e.g., SEM) is sufficient.

A higher degree of robustness, which is necessary for a serial application, could be achieved through the targeted integration of a certain variance into the training data according to a previously defined parameter space. Correspondingly more correlative data would be required for this. For a practical application, it also seems reasonable at this point to add augmentation routines optimized for the respective use case. This would most likely result in higher robustness. The data presented in this paper were collected under optimal and reproducible laboratory conditions in order to present the developed methodology as an example.

The unique characteristic of this work is that it unites separate approaches and combines them in a common workflow. Semantic segmentation of meaningful EBSD data was not used to quantify structures directly but to use the results specifically to achieve the greatest possible simplification and increase the efficiency of the underlying task. Thus, the primary requirement for this workflow was a high degree of objectivity combined with good reproducibility, as well as efficient applicability of the final segmentation model based on simple LOM images. It was possible to achieve an evaluability of highly complex structures where the established segmentation approaches reached their limits, especially due to the creation of a consistent ground truth.

Author Contributions: Conceptualization: B.-I.B., M.M. and D.B.; data curation: B.-I.B.; formal analysis: B.-I.B.; investigation: B.-I.B.; methodology: B.-I.B.; project administration: D.B., T.S. and F.M.; supervision: M.M., D.B., T.S. and F.M.; validation: B.-I.B.; visualization: B.-I.B.; writing—original draft: B.-I.B. and M.M.; writing—review and editing: B.-I.B., M.M., M.S., D.B., T.S. and F.M. All authors have read and agreed to the published version of the manuscript.

References

1. Nikas, D.; Zhang, X.; Ahlström, J. Evaluation of local strength via microstructural quantification in a pearlitic rail steel deformed by simultaneous compression and torsion. *Mater. Sci. Eng. A* **2018**, *737*, 341–347. [CrossRef]
2. Brandon, D.; Kaplan, W.D. *Microstructural Characterization of Materials*; John Wiley & Sons: Hoboken, NJ, USA, 2013.
3. van den Berg, F.; Kok, P.; Yang, H.; Aarnts, M.; Vink, J.J.; Beugeling, W.; Meilland, P.; Kebe, T.; Stolzenberg, M.; Krix, D.; et al. In-line characterisation of microstructure and mechanical properties in the manufacturing of steel strip for the purpose of product uniformity control. In Proceedings of the 19th World Conference on Non-Destructive Testing WCNDT, Munich, Germany, 13–17 June 2016.
4. Louthan, M.R., Jr. Optical Metallography. In *Materials Characterization*; ASM International: Almere, The Netherlands, 1986. [CrossRef]
5. Girault, E.; Jacques, P.; Harlet, P.; Mols, K.; Van Humbeeck, J.; Aernoudt, E.; Delannay, F. Metallographic Methods for Revealing the Multiphase Microstructure of TRIP-Assisted Steels. *Mater. Charact.* **1998**, *40*, 111–118. [CrossRef]
6. Samuels, L.E. *Light Microscopy of Carbon Steels*; ASM International: Almere, The Netherlands, 1999.
7. Britz, D.; Hegetschweiler, A.; Roberts, M.; Mücklich, F. Reproducible Surface Contrasting and Orientation Correlation of Low-Carbon Steels by Time-Resolved Beraha Color Etching. *Mater. Perform. Charact.* **2016**, *5*, 20160067. [CrossRef]
8. Campbell, A.; Murray, P.; Yakushina, E.; Marshall, S.; Ion, W. New methods for automatic quantification of microstructural features using digital image processing. *Mater. Des.* **2018**, *141*, 395–406. [CrossRef]
9. Collins, P.C.; Welk, B.; Searles, T.; Tiley, J.; Russ, J.C.; Fraser, H.L. Development of methods for the quantification of microstructural features in α+β-processed α/β titanium alloys. *Mater. Sci. Eng. A* **2009**, *508*, 174–182. [CrossRef]
10. Inkson, B.J. 2—Scanning electron microscopy (SEM) and transmission electron microscopy (TEM) for materials characterization. In *Materials Characterization Using Nondestructive Evaluation (NDE) Methods*; Hübschen, G., Altpeter, I., Tschuncky, R., Herrmann, H.-G., Eds.; Woodhead Publishing: Cambridge, UK, 2016; pp. 17–43. [CrossRef]
11. Zhu, K.; Barbier, D.; Iung, T. Characterization and quantification methods of complex BCC matrix microstructures in advanced high strength steels. *J. Mater. Sci.* **2013**, *48*, 413–423. [CrossRef]
12. Carneiro, Í.; Simões, S. Recent Advances in EBSD Characterization of Metals. *Metals* **2020**, *10*, 1097. [CrossRef]
13. Han, B.; Wan, W.; Sun, D.; Dong, C.; Zhao, L.; Wang, H. A deep learning-based method for segmentation and quantitative characterization of microstructures in weathering steel from sequential scanning electron microscope images. *J. Iron Steel Res. Int.* **2022**, *29*, 836–845. [CrossRef]
14. Szabó, P.J.; Bonyár, A. Effect of grain orientation on chemical etching. *Micron* **2012**, *43*, 349–351. [CrossRef]
15. Aarnts, M.P.; Rijkenberg, R.A.; Twisk, F.A.; Wilcox, D. *Microstructural Quantification of Multi-Phase Steels (Micro-Quant)*; European Comission: Luxembourg, 2011; Volume 10, p. 83656.
16. Zajac, S.; Schwinn, V.; Tacke, K.H. Characterisation and quantification of complex bainitic microstructures in high and ultra-high strength linepipe steels. *Mater. Sci. Forum* **2005**, *500–501*, 387–394.
17. Mayo, U.; Isasti, N.; Jorge-Badiola, D.; Rodriguez-Ibabe, J.M.; Uranga, P. An EBSD-based methodology for the characterization of intercritically deformed low carbon steel. *Mater. Charact.* **2019**, *147*, 31–42. [CrossRef]
18. Müller, M.; Britz, D.; Mücklich, F. Scale-bridging Microstructural Analysis—A Correlative Approach to Microstructure Quantification Combining Microscopic Images and EBSD Data. *Pract. Metallogr.* **2021**, *58*, 408–426. [CrossRef]
19. Holm, E.A.; Cohn, R.; Gao, N.; Kitahara, A.R.; Matson, T.P.; Lei, B.; Yarasi, S.R. Overview: Computer Vision and Machine Learning for Microstructural Characterization and Analysis. *Metall. Mater. Trans. A* **2020**, *51*, 5985–5999. [CrossRef]
20. Müller, M.; Stiefel, M.; Bachmann, B.-I.; Britz, D.; Mücklich, F. Overview: Machine Learning for Segmentation and Classification of Complex Steel Microstructures. *Metals* **2024**, *14*, 553. [CrossRef]
21. Stuckner, J.; Harder, B.; Smith, T.M. Microstructure segmentation with deep learning encoders pre-trained on a large microscopy dataset. *NPJ Comput. Mater.* **2022**, *8*, 200. [CrossRef]
22. Durmaz, A.R.; Müller, M.; Lei, B.; Thomas, A.; Britz, D.; Holm, E.A.; Eberl, C.; Mücklich, F.; Gumbsch, P. A deep learning approach for complex microstructure inference. *Nat. Commun.* **2021**, *12*, 6272. [CrossRef]
23. DeCost, B.L.; Francis, T.; Holm, E.A. Exploring the microstructure manifold: Image texture representations applied to ultrahigh carbon steel microstructures. *Acta Mater.* **2017**, *133*, 30–40. [CrossRef]
24. DeCost, B.L.; Lei, B.; Francis, T.; Holm, E.A. High throughput quantitative metallography for complex microstructures using deep learning: A case study in ultrahigh carbon steel. *Microsc. Microanal.* **2019**, *25*, 21–29. [CrossRef]

25. Decost, B.L.; Holm, E.A. A computer vision approach for automated analysis and classification of microstructural image data. *Comput. Mater. Sci.* **2015**, *110*, 126–133. [CrossRef]

26. DeCost, B.L.; Jain, H.; Rollett, A.D.; Holm, E.A. Computer Vision and Machine Learning for Autonomous Characterization of AM Powder Feedstocks. *JOM* **2017**, *69*, 456–465. [CrossRef]

27. Kopper, A.; Karkare, R.; Paffenroth, R.C.; Apelian, D. Model selection and evaluation for machine learning: Deep learning in materials processing. *Integr. Mater. Manuf. Innov.* **2020**, *9*, 287–300. [CrossRef]

28. Jacobs, R. Deep learning object detection in materials science: Current state and future directions. *Comput. Mater. Sci.* **2022**, *211*, 111527. [CrossRef]

29. Müller, M.; Britz, D.; Mücklich, F. Machine Learning for Microstructure Classification: How to Assign the Ground Truth in the Most Objective Way. *AMP Tech. Artic.* **2021**, *179*, 16–21. [CrossRef]

30. Bachmann, B.I.; Müller, M.; Britz, D.; Durmaz, A.R.; Ackermann, M.; Shchyglo, O.; Staudt, T.; Mücklich, F. Efficient reconstruction of prior austenite grains in steel from etched light optical micrographs using deep learning and annotations from correlative microscopy. *Front. Mater.* **2022**, *9*, 1033505. [CrossRef]

31. Laub, M.; Bachmann, B.I.; Detemple, E.; Scherff, F.; Staudt, T.; Müller, M.; Britz, D.; Mücklich, F.; Motz, C. Determination of grain size distribution of parental austenite grains through a combination of a modified contrasting method and machine learning. *Pract. Metallogr.* **2022**, *60*, 4–36. [CrossRef]

32. Bachmann, B.-I.; Müller, M.; Britz, D.; Staudt, T.; Mücklich, F. Reproducible Quantification of the Microstructure of Complex Quenched and Quenched and Tempered Steels Using Modern Methods of Machine Learning. *Metals* **2023**, *13*, 1395. [CrossRef]

33. Breumier, S.; Ostormujof, T.M.; Frincu, B.; Gey, N.; Couturier, A.; Loukachenko, N.; Aba-perea, P.E.; Germain, L. Leveraging EBSD data by deep learning for bainite, ferrite and martensite segmentation. *Mater. Charact.* **2022**, *186*, 111805. [CrossRef]

34. Shen, C.; Wang, C.; Huang, M.; Xu, N.; van der Zwaag, S.; Xu, W. A generic high-throughput microstructure classification and quantification method for regular SEM images of complex steel microstructures combining EBSD labeling and deep learning. *J. Mater. Sci. Technol.* **2021**, *93*, 191–204. [CrossRef]

35. Nowell, M.M.; Witt, R.A.; True, B.W. EBSD Sample Preparation: Techniques, Tips, and Tricks. *Micros Today* **2005**, *13*, 44–49. [CrossRef]

36. Gazder, A.A.; Al-Harbi, F.; Spanke, H.T.; Mitchell, D.R.G.; Pereloma, E.V. A correlative approach to segmenting phases and ferrite morphologies in transformation-induced plasticity steel using electron back-scattering diffraction and energy dispersive X-ray spectroscopy. *Ultramicroscopy* **2014**, *147*, 114–132. [CrossRef]

37. Wright, S.I.; Nowell, M.M. EBSD image quality mapping. *Microsc. Microanal.* **2006**, *12*, 72–84. [CrossRef] [PubMed]

38. Wright, S.I.; Nowell, M.M.; Field, D.P. A review of strain analysis using electron backscatter diffraction. *Microsc. Microanal.* **2011**, *17*, 316–329. [CrossRef]

39. Mainprice, D.; Bachmann, F.; Hielscher, R.; Schaeben, H. *Descriptive Tools for the Analysis of Texture Projects with Large Datasets Using MTEX: Strength, Symmetry and Components*; Geological Society: London, UK, 2015; Volume 409, pp. 251–271.

40. Schneider, C.A.; Rasband, W.S.; Eliceiri, K.W. NIH Image to ImageJ: 25 years of image analysis. *Nat. Methods* **2012**, *9*, 671–675. [CrossRef]

41. Sorzano, C.O.S.; Thévenaz, P.; Unser, M. Elastic registration of biological images using vector-spline regularization. *IEEE Trans. Biomed. Eng.* **2005**, *52*, 652–663. [CrossRef] [PubMed]

42. Waterschoot, T.; Kestens, L.; De Cooman, B.C. Hot Rolling Texture Development in CMnCrSi Dual-Phase Steels. *Metall. Mater. Trans. A* **2002**, *33*, 1091–1102. [CrossRef]

43. De Vincentis, N.S.; Field, D.P. Factors affecting Confidence Index in EBSD analysis. *Ultramicroscopy* **2021**, *225*, 113269. [CrossRef]

44. Ryde, L. Application of EBSD to analysis of microstructures in commercial steels. *Mater. Sci. Technol.* **2006**, *22*, 1297–1306. [CrossRef]

45. Li, X.; Ramazani, A.; Prahl, U.; Bleck, W. Quantification of complex-phase steel microstructure by using combined EBSD and EPMA measurements. *Mater. Charact.* **2018**, *142*, 179–186. [CrossRef]

46. Zaefferer, S.; Romano, P.; Friedel, F. EBSD as a tool to identify and quantify bainite and ferrite in low-alloyed Al-TRIP steels. *J. Microsc.* **2008**, *230*, 499–508. [CrossRef]

47. Radwański, K. Structural characterization of low-carbon multiphase steels merging advanced research methods with light optical microscopy. *Arch. Civ. Mech. Eng.* **2016**, *16*, 282–293. [CrossRef]

48. Man, O.; Pantělejev, L.; Man, O.; Pešina, Z. EBSD analysis of phase compositions of TRIP steel on various strain levels. *Mater. Eng.* **2009**, *16*, 15–21.

49. Na, S.-H.; Seol, J.-B.; Jafari, M.; Park, C.-G. A Correlative Approach for Identifying Complex Phases by Electron Backscatter Diffraction and Transmission Electron Microscopy. *Appl. Microsc.* **2017**, *47*, 43–49. [CrossRef]

50. Wu, J.; Wray, P.J.; Garcia, C.I.; Hua, M.; DeArdo, A.J. Image quality analysis: A new method of characterizing microstructures. *ISIJ Int.* **2005**, *45*, 254–262. [CrossRef]

51. Kang, J.Y.; Do, H.K.; Baik, S.I.; Ahn, T.H.; Kim, Y.W.; Han, H.N.; Oh, K.H.; Lee, H.C.; Han, S.H. Phase analysis of steels by grain-averaged EBSD functions. *ISIJ Int.* **2011**, *51*, 130–136. [CrossRef]

52. Mo, Y.; Wu, Y.; Yang, X.; Liu, F.; Liao, Y. Review the state-of-the-art technologies of semantic segmentation based on deep learning. *Neurocomputing* **2022**, *493*, 626–646. [CrossRef]

53. Ostormujof, T.M.; Purohit, R.R.P.P.R.; Breumier, S.; Gey, N.; Salib, M.; Germain, L. Deep Learning for automated phase segmentation in EBSD maps. A case study in Dual Phase steel microstructures. *Mater. Charact.* **2022**, *184*, 111638. [CrossRef]
54. Na, J.; Kim, S.J.; Kim, H.; Kang, S.H.; Lee, S. A unified microstructure segmentation approach via human-in-the-loop machine learning. *Acta Mater.* **2023**, *255*, 119086. [CrossRef]
55. Jadon, S. A survey of loss functions for semantic segmentation. In Proceedings of the 2020 IEEE Conference on Computational Intelligence in Bioinformatics and Computational Biology (CIBCB), Via del Mar, Chile, 27–29 October 2020. [CrossRef]
56. Qubvel/Segmentation_Models: Segmentation Models with Pretrained Backbones. Keras and TensorFlow Keras. Available online: https://github.com/qubvel/segmentation_models (accessed on 9 August 2022).
57. Szegedy, C.; Vanhoucke, V.; Ioffe, S.; Shlens, J.; Wojna, Z. Rethinking the inception architecture for computer vision. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016; pp. 2818–2826.
58. Rodríguez, C.G.; Vitrià, J.; Mora, O. Uncertainty-based human-in-the-loop deep learning for land cover segmentation. *Remote Sens.* **2020**, *12*, 3836. [CrossRef]