



Grain size analysis in permanent magnets from Kerr microscopy images using machine learning techniques

Amit K. Choudhary^{a,b,*}, Andreas Jansche^a, Tvrtko Grubesa^{a,b}, Florian Trier^a, Dagmar Goll^a, Timo Bernthaler^a, Gerhard Schneider^a

^a Materials Research Institute, Aalen, Germany

^b Karlsruhe Institute of Technology, Karlsruhe, Germany

ARTICLE INFO

Keywords:

Machine learning
Kerr microscopy
Permanent magnets
Grain analysis
Semantic segmentation
Micromagnetic domains
Quantitative microstructure analysis

ABSTRACT

Understanding the relationships between composition, structure, processing and properties helps in the development of improved materials for known applications as well as for new applications. Materials scientists, chemists and physicists have researched these relationships for many years, until the recent past, by characterizing the bulk properties of functional materials and describing them with theoretical models.

Magnets are widely used in electric vehicles (EV), hybrid electric vehicles (HEV), data storage, power generation and transmission, sensors etc. The search for novel magnetic phases requires an efficient quantitative microstructure analysis of microstructural information like phases, grain distribution and micromagnetic structural information like domain patterns, and correlating the information with intrinsic magnetic parameters of magnet samples. The information out of micromagnetic domains helps in obtaining the optimized microstructures in magnets that have good intrinsic magnetic properties.

This paper is aimed at introducing the use of a traditional machine learning (ML) model with a higher dimensional feature set and a deep learning (DL) model to classify various regions in sintered NdFeB magnets based on Kerr-microscopy images. The obtained results are compared against reference data, which is generated manually by subject experts. Additionally, the results were compared against the approach for grain analysis, which is based on the electron backscatter diffraction (EBSD) technique. Further, the challenges faced by the traditional machine learning model for classifying microstructures in Kerr micrographs are discussed.

1. Introduction on hard-magnetic materials and their microstructure

One of the main factors for the improved energy efficiency of EVs and HEVs is the usage of permanent magnets in synchronous motors (high coercivity rare earth sintered magnets) [1,2]. The Neodymium Iron Boron (NdFeB) magnets are one of the most commonly used magnets in motors for EVs and HEVs. NdFeB magnets have higher remanence, much higher coercivity and energy product, but lower Curie temperature than the alternatives such as Ferrite magnets, AlNiCo and Samarium Cobalt magnets [3]. To resist the demagnetization field during operation, the coercive force of the NdFeB magnets is improved by partially replacing neodymium with heavy rare earth (RE) elements like dysprosium or terbium, or by precise optimization of the microstructure [4]. However, these rare earth metals are very expensive and limited supply are available globally. Therefore, there is demand for permanent magnets

with reduced RE content inexpensive RE metals [5,6].

The microstructure of polycrystalline (Nd,RE)FeB-based permanent magnets (RE: substitutional rare-earth elements, e.g. Ce, Pr, Dy) is made up of independent μm -sized (Nd,RE)₂Fe₁₄B-grains with high uniaxial magneto-crystalline anisotropy. During manufacturing the (Nd,RE)FeB-particles are aligned with their magneto-crystalline anisotropy axis parallel to an external magnetic field which produces a strong preferential orientation distribution in this direction in order to reach the maximum possible remanence of the magnetic material. Ideally, the aligned grains are separated by a thin grain boundary phase, which forms upon liquid phase sintering and in solidified state minimizes the magnetic influence of one grain on another. During the sintering process grain growth can occur in an undesirable manner and lead to abnormal grain growth. So, the monitoring of the grain size distribution is a part of quality assurance in magnet production. Sintered magnets can furthermore contain pores and rare earth oxide particles, which lead to

* Corresponding author at: Materials Research Institute, Aalen, Germany.

E-mail address: amit.kumar.choudhary@hs-aalen.de (A.K. Choudhary).

<https://doi.org/10.1016/j.matchar.2022.111790>

Received 2 September 2021; Received in revised form 2 February 2022; Accepted 9 February 2022

Available online 11 February 2022

1044-5803/© 2022 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

diminished volume fraction of the permanent magnetic phase and thus to lowered remanence.

In the demagnetized state, the net magnetization of a grain is zero. This state is characterized by the formation of regions of uniform but opposite magnetization within the grain, which are called magnetic domains. Domains form as a result of minimizing the total energy which arises from the magnetic dipole moments: the magnetostatic energy contribution is minimized by reducing the magnetic field outside the grain, so that the magnetic field can form a loop into a region of opposite magnetization. The energy required to form this domain is defined by the domain wall energy, which is then stored in the boundary of the two domains, the so-called Bloch wall. The equilibrium state is reached when both energy contributions are balanced out, which leads to a characteristic domain pattern of the material, depending on the anisotropy constant and the saturation magnetization. In high anisotropy (Nd, RE)₂Fe₁₄B-grains the magnetic domains are oriented along the magnetocrystalline anisotropy axis of the grain. In adjacent domains the magnetization changes direction from 0° to 180°.

Magnetic imaging techniques give one of the most appropriate view of the magnetic properties on a microscopic scale. The phase distribution, domain patterns, grain size distribution, regions of the same magnetization direction and their arrangement as a function of geometry and its properties can be studied using microscopy [7]. Fig. 1 shows the microstructure and micromagnetic structures (domains) of the anisotropic magnet sample in the demagnetized state using light microscope (LM) and Kerr microscope [8] using a correlative microscopy approach. The magnet sample was prepared through axial compaction

and sectioned at 90° (stripe domains) and 0° (closure domains) to observe different micromagnetic structures. Further, the domains appear differently depending upon the sample microsection with respect to the axis of preferential alignment. In stripe domains the patterns inside grains are aligned in orderly manner and for closure domain structures they appear to be randomly aligned. Each microscopy gives a different scale of information about the microstructure. The LM image gives a better view of phase distribution, while Kerr micrographs give information about the grain size distribution and domain patterns. Microstructural parameters also have a strong influence on the properties of the NdFeB magnets [7,9]. D. Goll et al. [10] explained the importance of quantitative microstructure analysis of domain patterns of hard magnetic phases to find intrinsic magnetic parameters. The saturation polarization from domain pattern contrast difference and anisotropy constant from the width of patterns was calculated using the domain images obtained from Kerr microscopy for the NdFeB magnet. Lukas et al. [11] showed the application of machine learning for the characterization of microstructure features in ideally structured Nd₂Fe₁₄B magnets and concluded that the position of the grain within the magnet is of high importance when compared to the grains near the top and bottom edges.

1.1. Magneto-optic Kerr effect

The range of domain observation techniques spans from optical and electron microscopy to magnetic force microscopy. Among these, the magneto-optic Kerr effect (MOKE) can visualize domains with a spatial

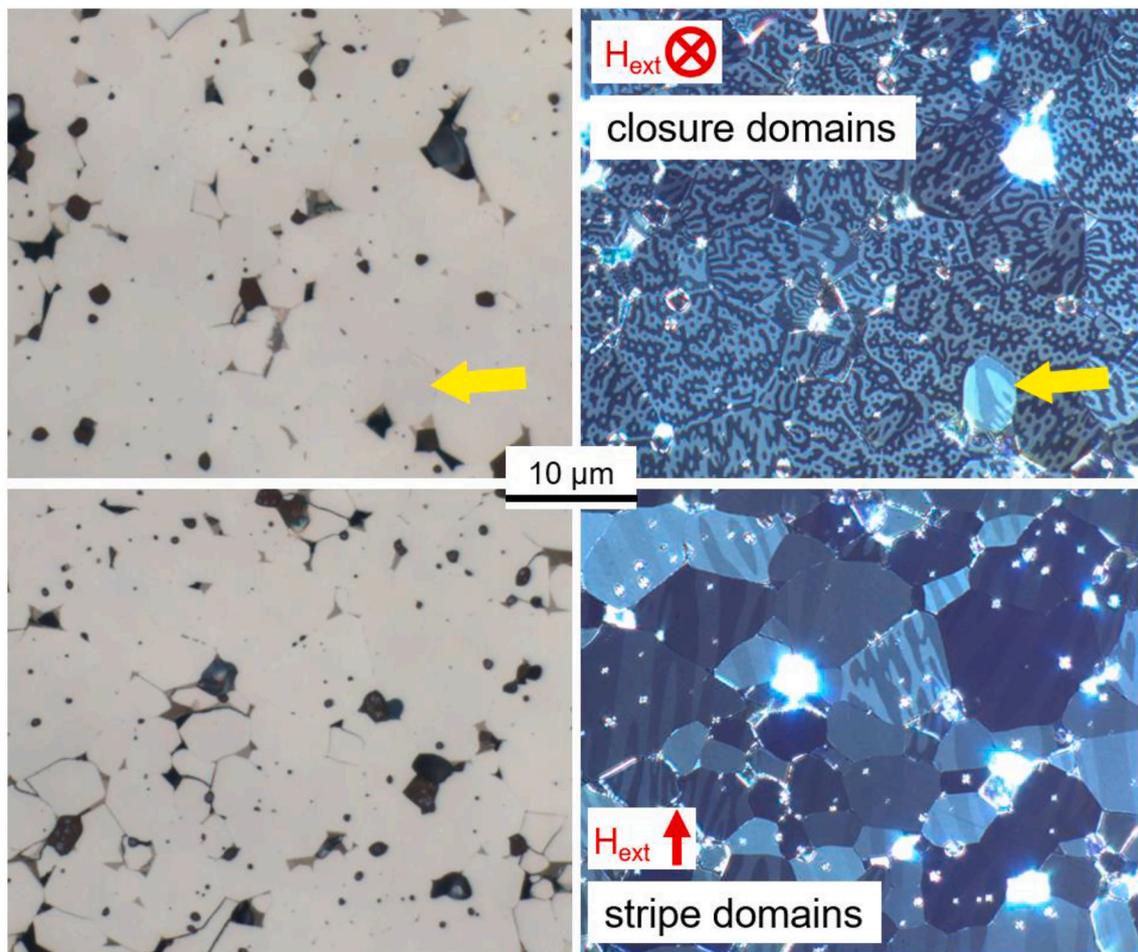


Fig. 1. Microstructure of the NdFeB magnet sample acquired using correlative microscopy showing different domain types: light optical microscopic micrograph and Kerr micrograph, 1000× magnification.

resolution of 0.1 μm , recording time of several hundred milliseconds to several seconds and information depth up to 20 nm using an optical microscope [7]. The Kerr effect describes the interaction between a beam of linearly polarized light and a ferromagnetic sample. Upon reflection of the beam of incident linear s-polarization becomes partly p-polarized, which corresponds to an elliptical polarization. The angle of the main axis of the polarization ellipse with respect to the incident linear polarization plane is the Kerr rotation angle, which is dependent of the saturation magnetization of the material. Depending on the geometry of incident light with respect to the direction of the magnetization axis, different MOKE can be distinguished: longitudinal, transversal and polar as described by Schäfer et al. [7]. In our experiments the polar MOKE setup was used, yielding the polar magnetization component of the domains.

One of the techniques for quantitative microstructure analysis of magnetic materials is using Kerr microscopy, which has the advantage of being relatively inexpensive, non-invasive and able to handle a large range of magnetic samples [12,13,7,8]. Fig. 2 shows the Kerr image of an anisotropic NdFeB sample sectioned at 60° to observe micromagnetic structures (domains) and grain boundaries highlighted in red color across the sample and the domain patterns of some of the grains highlighted in yellow color. This paper is focused on the application of ML models to get grain size distribution in permanent magnet samples from Kerr images and compare results against reference data obtained manually by subject experts.

2. Machine learning for microstructure analysis

Automated quantitative microstructure analysis is a widely accepted method for quality control in materials, especially for grain size analysis. The ability of machine learning models to overcome the challenge of developing complex multivariate functions by considering human expert knowledge has been one of the major factors in their success. Within the scope of this paper, a grain analysis has been developed and evaluated for magnet samples using a traditional machine learning approach, involving feature extraction and advanced deep learning models for Kerr microscopy applications.

Fig. 3 shows the workflow of the supervised machine learning model, in which a region of interest or expected output are to be provided to the training model as labels. A traditional machine learning approach involves a feature vector, which is a matrix containing extracted image features from the training dataset and the corresponding label. Feature

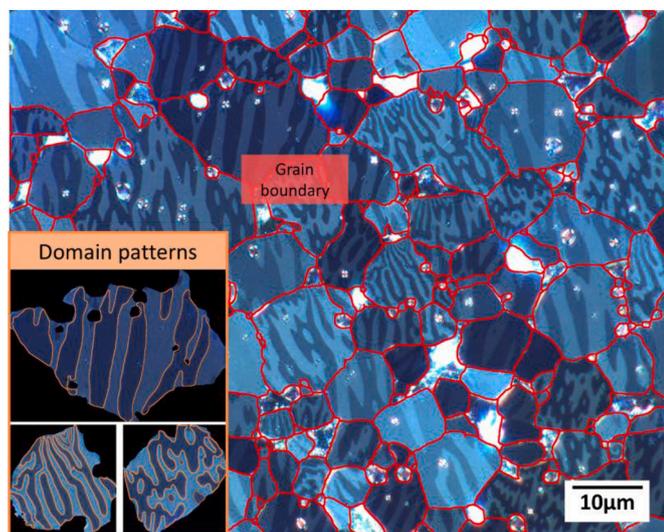


Fig. 2. Shows the Kerr microscopic image of a commercially available $\text{Nd}_2\text{Fe}_{14}\text{B}_1$ sinter magnet acquired under $1000\times$ magnification, with grains and domain patterns within the grains highlighted for reference.

extraction is one of the methods used to interpret the input data for machine learning tasks such that the information extracted is non-redundant and is expected to have the necessary information about input in reduced form. It helps the machine learning model with learning and generalization steps during the training phase. It allows us to represent images in a higher-dimensional space, which helps with pixel classification.

Based on the pre-defined learning task, features such as color, edge features, texture features, etc., can vary. Table 1 shows some of the widely used image features and the category to which they belong.

The image feature extraction is highly subjective based on the expected output from the model and on the complexity of the microstructure of the sample. The features extracted from input have a direct effect on model performance; therefore the user must know different image processing techniques in order to design a suitable feature vector that contains relevant information so that the model can find the relationship between the input and the expected output. The advanced machine learning approach, including deep learning techniques, would use the model to automatically extract high-level features from the input data. This reduces the need for domain expertise for manual feature extraction steps [14]. Such models have proven to be more accurate and robust when compared to traditional feature-based machine learning models [15]. However, deep learning models require a large dataset and high computational expenses for training the model.

In the case of permanent magnet samples, the appearance of the grains in Kerr images shows that the color difference between individual grains is very fine, as shown in Fig. 1, and therefore a higher dimensional feature vector is required to classify each grain. Further, the extraction of features from such a complex microstructure is a tedious task. In addition, generating the labels for Kerr images (grain boundaries) is time-consuming as it is performed manually by domain experts. Therefore, we would need a state-of-the-art DL model such as U-net, which would work on the small datasets and avoid tedious manual feature extraction steps [16].

U-net has a convolutional neural network (CNN) architecture developed for the semantic segmentation to perform with high accuracy, and requires a relatively small dataset for training the model. Fig. 4 shows that the network has a contracting path that follows the architecture of a general convolutional network and has multiple blocks of 3×3 convolutions followed by a rectified linear unit (ReLU) and a 2×2 max pooling with stride 2 for downsampling. The number of kernels (feature maps) doubles after each block, which helps the model learn complex structures effectively. The bottom-most layer connects the contraction and expansion layer with two 3×3 CNN layers followed by a 2×2 up convolutional layer. The crucial part lies in the expansion layer, which has several expansion blocks of 3×3 CNN layers and an upsampling 2×2 layer. After each block, the feature map is reduced by half to maintain the symmetry. Further, each time input is also appended by a feature map of the corresponding contraction layer to ensure that the features learned are used for reconstruction. In the final layer, a 1×1 convolutional layer is used to join each of the 64 component feature vectors to the desired labels [16].

3. State of the art

Over the years, the importance of grain size analysis has grown and different approaches have been developed for different use cases; these include traditional image analysis as well as ML-based approaches. There is not much literature available on approaches to automatically detect and analyze grain distribution from Kerr images of magnets, and to the best of our knowledge the approaches in this paper are the first work on this problem. However, there have been considerable efforts made in recent years to use machine learning approaches to detect phase boundaries in the microstructures of the materials.

Pusch et al. [9] developed a semi-automated approach to quantify domain structure using correlative imaging. The NdFeB sinter magnet

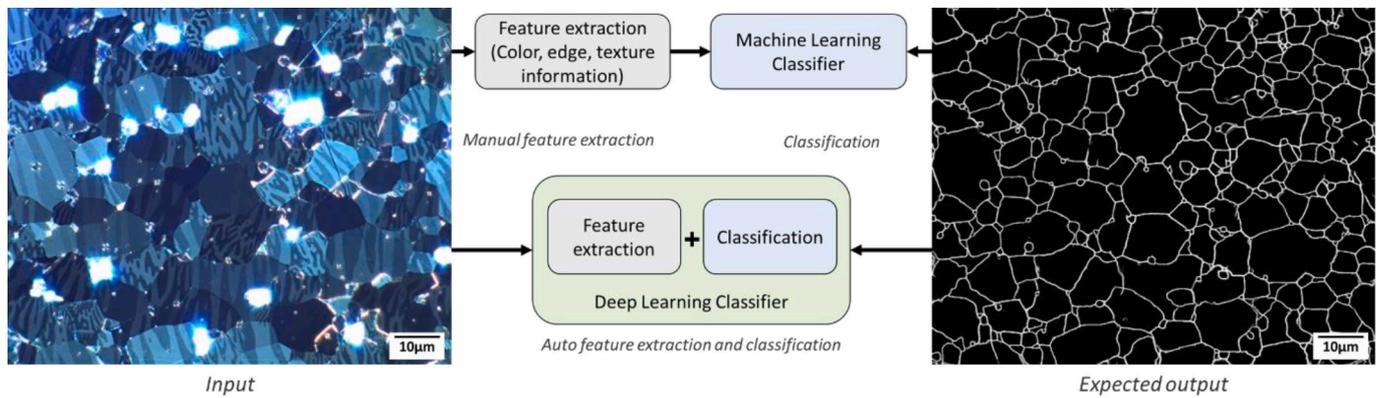


Fig. 3. Shows the workflow for training and evaluation of the supervised machine learning model. The workflow of traditional machine learning involves a data preprocessing step to extract relevant image features manually, and expert labeling to the learning model (top). The deep learning approach involves auto feature extraction and is highly dependent on the size of the training dataset (bottom).

Table 1

Shows commonly used features that can be extracted from an image and the associated image information, which can be used for training the machine learning models.

Features	Image information
RGB (Red, Green, Blue) channels	Color
HSV (Hue, Saturation, Value) channels	Color
Gray value	Color
GLCM (Gray level co-occurrence matrix)	Texture
Structure tensor	Texture
Histogram of gradients	Texture
Median filter	Texture
Bilateral filter	Edge
Sobel filter	Edge
Difference of Gaussians	Edge

sample was used to acquire a correlative image under a polarized light microscope and the optical microscope to get grain boundaries. The two images were mapped onto each other to find the grain boundaries; using

the maximum threshold and second maximum threshold from the histogram of each grain, domain patterns were extracted as shown in Fig. 5. The extracted domains were then classified into six groups for analysis. This approach requires the sample to be prepared with special care to make sure grain boundaries are visible in LM images, and careful acquisition of correlative images and human intervention if the grain boundaries in the two images do not match.

The U-net architecture was used for grain boundary detection by Furat et al. [17] in his work, to obtain grain boundaries of the time-resolved computer tomography data from the AlCu sample. The images were processed using 2D U-net, multichannel 2D U-net and 3D U-net to obtain the comparative segmentation results. The 3D U-net performed better than the hand-labeled ground truth. The approach used in this paper highlights the importance of precise labeling of the dataset to ensure better predictions.

Neuman et al. [18] proposed a hybrid algorithm combining a random forest classifier and watershed algorithm to trinarize the image data of Ibuprofen tablets to detect pores, API and MCC. The random forest classifier was trained on the higher dimensional feature set to

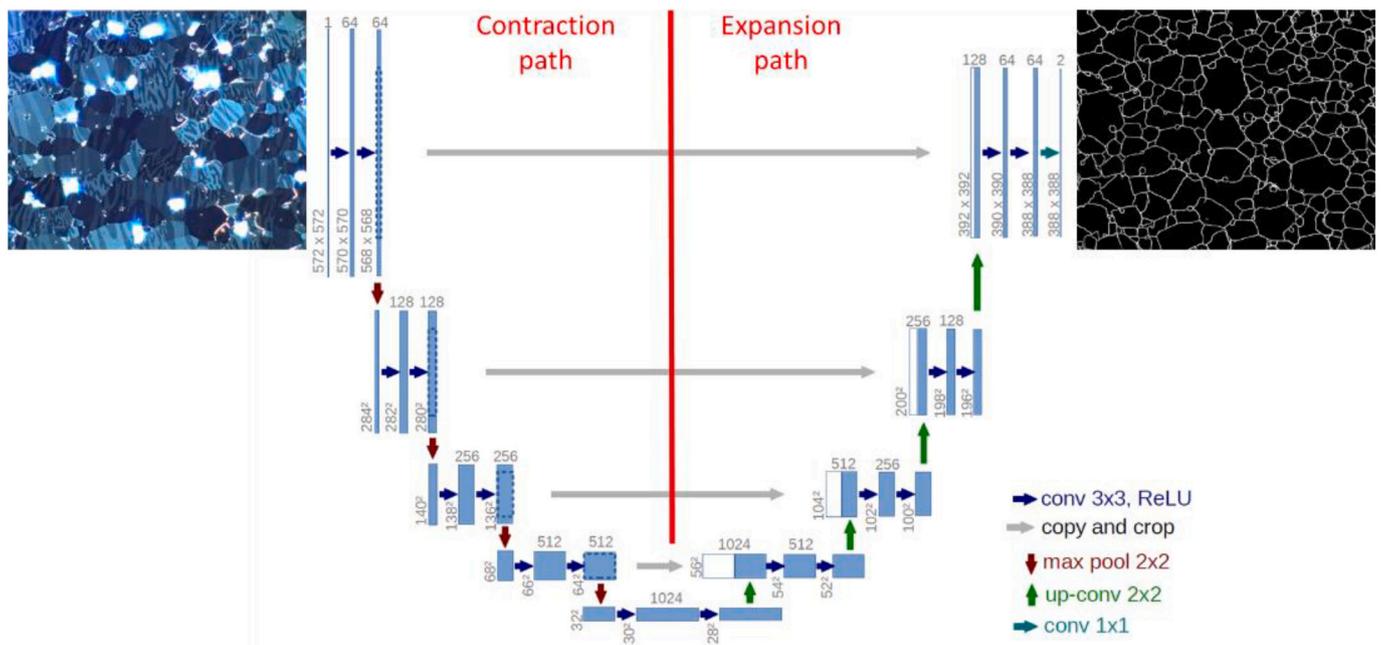


Fig. 4. Shows the architecture of the U-net algorithm. The architecture has a contracting path on the left side and an expansion path on the right side. The blue boxes represent multi-channel feature maps and the white boxes are copied feature maps. Different-colored arrows represent different operations as shown at the bottom right [16].

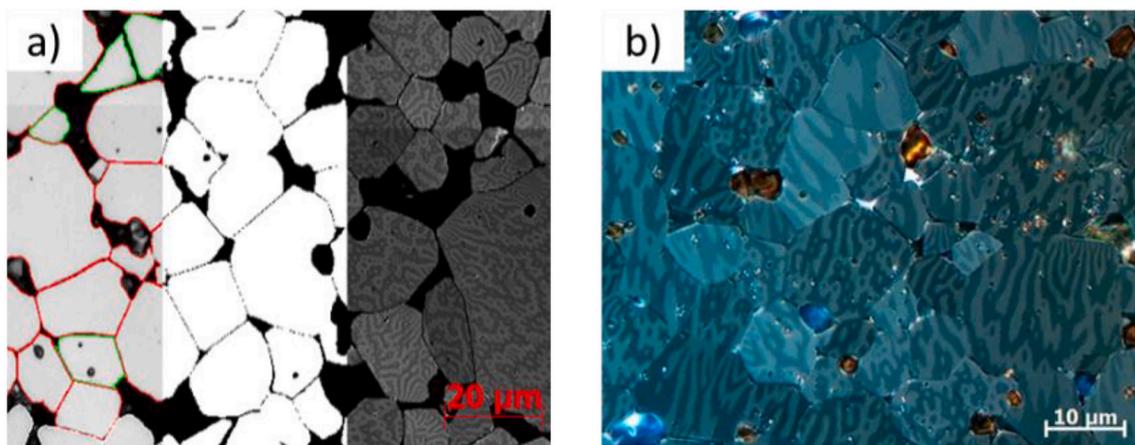


Fig. 5. Shows (a) a light microscopic image of the NdFeB magnet sample with grain boundaries obtained after threshold-based image segmentation, which is mapped onto a correlative Kerr image to extract grains and (b) a sample Kerr micrograph image [9].

classify the three regions of interest. The author compared the image trinarization performance using watershed, random forest classifier and hybrid models. The hybrid algorithm proved to be better overall as it optimized the accuracy of the task. This method has some limitations, as the watershed algorithm is not stable and often requires the user to pre-define the markers in order to obtain satisfactory results. The feature-based machine learning approach was also adopted by Jiang et al. [19] in his work, which is based on superpixel extraction and then creating a new higher-dimensional feature set that includes extraction of color, texture and edge information from each superpixel for grain segmentation in thin-section images of rocks. Under a polarized light microscope, multiple images were acquired at different angles; next, simple linear iterative clustering (SLIC) was adopted for over-segmenting each image based on the pixel intensity criteria. This approach resulted in an F1 score of 0.87 when validated for the maximum-intensity image.

In a broader context, the task of grain boundary detection in the material samples is similar to edge detection in natural images or biomedical images. There have been many state-of-the-art DL models used for edge detection, such as DeepContour [20], DeepEdge [21], Sketch Tokens [22] and Holistically Nested Edge Detection (HED) [23]. These models were evaluated with the Berkeley segmentation dataset (BSDS) and benchmark [24]. The dataset has 200 training, 100 validation, and 200 testing images. The framework of the HED model achieved better results when compared to recent CNN-based edge detectors. The HED model achieved a fixed contour threshold (ODS) of 0.78 and was close to the human ODS of 0.80. These models were trained and performed fairly well on natural images, but fail to detect grain boundaries or required edges in materials science images such as magnet Kerr

micrographs.

Fig. 6 shows the result of the HED model on one of the sample magnet images acquired using the Kerr microscopy. The HED model performed well in detecting the grain boundaries in some regions within the image but failed to classify the grain boundaries from domain patterns. Therefore, there is a need to have an edge detection model customized to work on magnet samples when these are analyzed using Kerr microscopy.

The other standard approach for grain analysis is the EBSD technique as described by Hohns et al. [25]. EBSD has been a highly effective and useful tool for the grain analysis of various materials [26]. The EBSD approach has many advantages such as analyzing grain size distribution, grain texture analysis and grain orientation, etc. [26–28]. However, if the grain size or the grain orientation has to be analyzed over a larger area, e.g. to get the distribution of size or orientation in a complete component, EBSD is too time-consuming and not suitable.

Therefore, an attempt has been made to apply ML techniques along with computer vision steps to get grain size distribution directly from light microscopic images of magnetic materials, with higher accuracy and less processing time than EBSD and other manual approach.

4. Experiments

4.1. Materials and dataset

The dataset consists of Kerr images of four different commercially available NdFeB sintered permanent magnets of different alloys in demagnetized state acquired at $1000\times$ magnification and the Kerr effect was visualized with a polarizer-analyzer pair with 88° angle between the

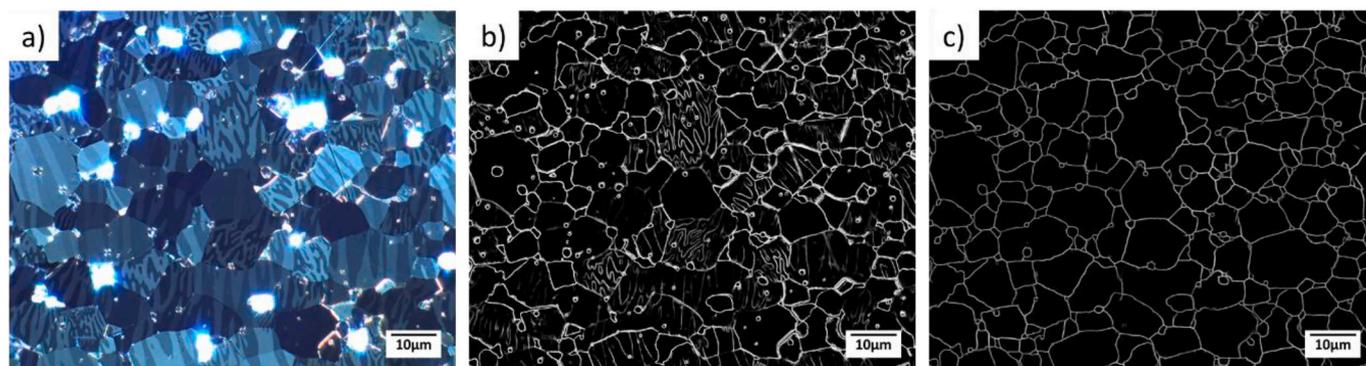


Fig. 6. Shows an image of an NdFeB sintered magnet ($1000\times$) (a) that has been processed to predict grain boundaries from the HED model trained on the BSDS dataset (b) and a U-net model trained on the magnet dataset (c).

planes of polarization. The samples NdFeB-A and NdFeB-B are anisotropic produced by axial compaction. The remanence (B_R) and coercivity (H_{cJ}) of NdFeB-A is 1.3 T and 1355 kA/m whereas for NdFeB-B it is 1.08 T and 2625 kA/m respectively. The samples NdFeB-C and NdFeB-D are also anisotropic in nature produced by axial compaction but their B_R and H_{cJ} is not known and were considered for experiments to study the robustness of the learning models. The images of NdFeB-C and NdFeB-D samples were acquired using a Zeiss AxioImager.M2 with exposure time of 10,000 ms under analyzer reflector light. The images of NdFeB-A and NdFeB-B samples were acquired using a Zeiss AxioImager.Z2 with exposure time of 1800 ms under polarized reflector light. The sample consists of grains which are primarily oriented along the magnetocrystalline anisotropy axis of the grain. This would mean that the samples have high amount of grains with stripe domain structure. The acquired images were labeled manually to train and evaluate the performance of models. The manual approach for labeling the images required a user with subjective knowledge to trace the grain boundaries on the Kerr images. The sample hand-labeled image for the corresponding Kerr image is shown in Fig. 7.

The labeled images obtained from the manual approach are later processed using various image processing steps to get grain boundaries, which are visible as black pixels, and the red pixels indicate the direction of the domain patterns of the grains with strip domains (seen in Fig. 7). The images were acquired at $1000\times$ magnification using an oil immersion objective. As seen in Fig. 8, the area of the samples used for experiments is large and therefore the acquired images are split into tiles (1525×1128 pixels) to reduce computational expenses when processing the images for model training. Table 2 gives a brief description of the samples and the number of images used for the experiment. The dataset was split into three sets that include training images, validation images and testing images. The training and validation images are the ones that the model has seen during the model training phases (multiple iterations) and the testing images are those that the model has not seen before; these are tested after the model has been successfully trained to achieve significant performance. Further, for the sample NdFeB-A and NdFeB-B EBSD maps were also acquired using a Zeiss Sigma 300 VP, EDAX scanning electron microscope. However, it should be noted that the acquired EBSD map does not perfectly correlate to the Kerr images. The EBSD maps tend to remove small grains during the post-processing steps to clean the noisy data received from the scanning electron microscope. Therefore, as a reference or ground truth to measure the performance of the machine learning models hand-labeled data was considered instead of EBSD.

4.2. Machine learning and deep learning models for Kerr images

For the experiments, two different approaches for grain detection

and analysis were considered. The first approach involved using the traditional machine learning workflow, which involves feature-engineering steps. Feature-engineering involves the extraction of the useful information from the input images to be processed by the ML model. The other approach involves using a deep learning model that performs feature extraction on its own to learn useful information to predict the grain boundaries from the input images. The difference in both approaches is illustrated in Fig. 9. As a common preprocessing step for both approaches, the preparation artifacts, pores and oxides are removed from the images using the threshold-based image-processing method.

The feature-engineering step involves creating a 39-dimensional feature vector that includes 6 color features (RGB and HSV channels), 15 texture features (median filter and structure tensor) and 18 edge features (bilateral filter, difference of Gaussians and Sobel filter). All of these features are extracted for different levels of Gaussian blur (sigma: 1, 2 and 4). The workflow involves 39 feature maps (one for each feature extracted) and therefore we have 39 corresponding feature values for each pixel in the input image. The workflow is supervised learning and thus involves expected outputs as labels to train the model. Here, we had 20 different classes or types of grains to be classified by a manually optimized random forest classifier [29]. Random forest classifier is an ensemble tree-based learning algorithm that decides the class of the input based on votes from multiple decision trees. This algorithm has the advantage of running efficiently on large datasets and gives feedback to the user on important features contributing to the learning process. In addition, this algorithm is prone to overfitting if hyper-parameters are not fine-tuned effectively. The hyperparameters of the optimized classifier used here have 100 trees, Gini function to measure the quality of the split, a maximum depth of 100 for the tree, and out-of-bag samples to generalize the accuracy. The trained model when tested on the validation set achieved an F1 score of 0.93. The segmented image from a trained random forest classifier needs post-processing to remove the artifacts and improve the grain boundary detection. Using the superpixel clustering approach, the artifacts are removed. The importance of training the random forest classifier on a higher-dimensional feature set can be visualized from Fig. 9. The ML model trained on color features only was unable to classify the grains and therefore additional image features were extracted to improve the model performance.

The deep learning model used for the experiment here is U-net, whose architecture works on the small training dataset when the architecture is compared to other CNNs. It is a supervised learning approach, and the datasets used for training, testing and validation are the same for both random forest classifier and U-net. The architecture of the model is described in Section 2. The architecture is synonymous with encoder-decoder; feature extraction is carried out in the contracting path, and the expanding path does the localization to retain the spatial

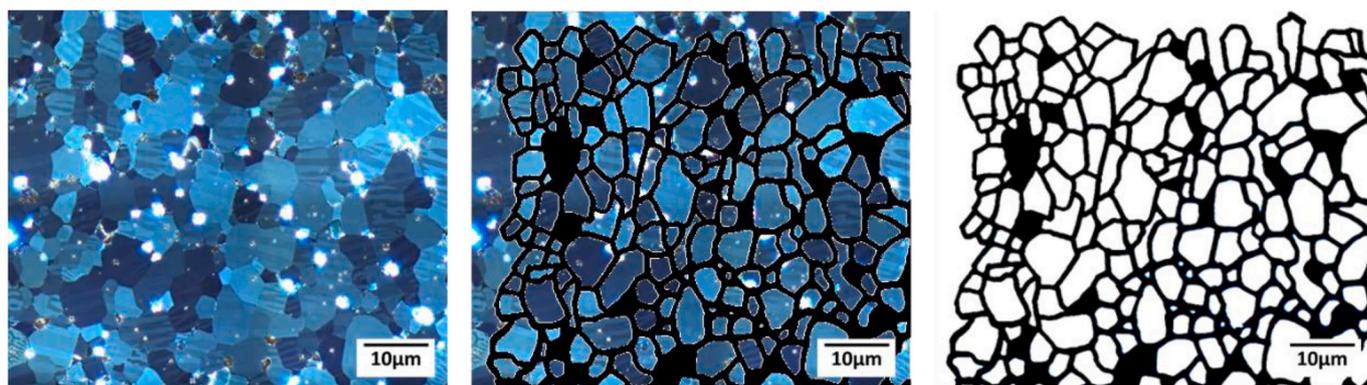


Fig. 7. Shows the workflow of grain boundaries detected in NdFeB sinter magnet sample using a manual approach from Kerr micrographs. The grain boundaries are traced using a black marker. This hand-traced image is post-processed using threshold based image segmentation, erosion and calculating the morphological parameters to get the grain distribution curve. (left to right).

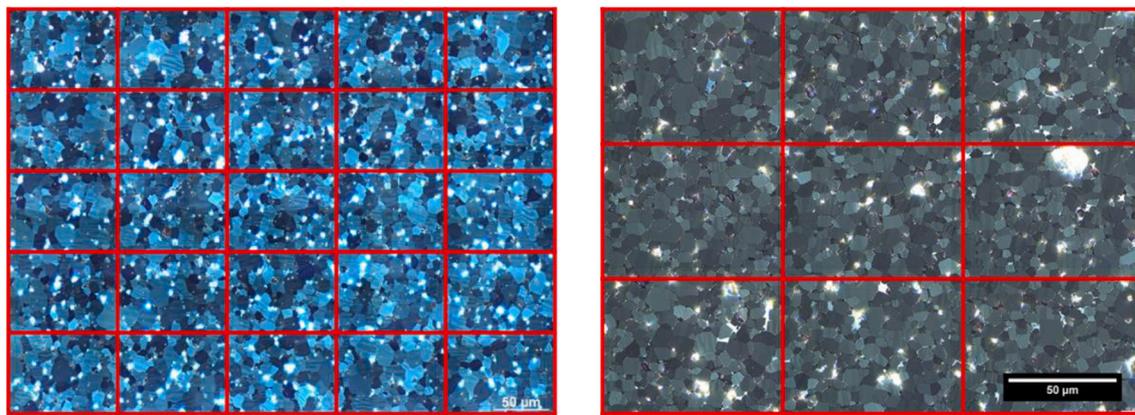


Fig. 8. Shows the Kerr micrograph (1000 \times) of NdFeB-A and NdFeB-D split into tiles based on the area of the sample. The area of the NdFeB-A sample is 457 \times 338 μm^2 and NdFeB-D has an area of 233 \times 174 μm^2 . The tile images are sub-divided into training, validation and test datasets.

Table 2

Shows the overview of the different magnet samples used for training, validation and testing of the U-net model. The acquired Kerr micrograph of each sample is a tile image with multiple tiles as mentioned in column 2.

Sample ID	Kerr image (as tiles)	EBS D map	Hand labeled data	Training dataset	Validation dataset	Testing dataset
NdFeB-A	25	Yes	Yes	10	12	03
NdFeB-B	15	Yes	Yes	00	00	15
NdFeB-C	09	No	Yes	00	00	09
NdFeB-D	09	No	Yes	04	03	02

NdFeB-A and NdFeB-B: Area of each tile 92 \times 67 μm^2 .

NdFeB-c and NdFeB-D: Area of each tile 78 \times 58 μm^2 .

information despite downsampling and max-pooling performed during the contracting path. We used Adam as optimizer, sigmoid was used as the activation function, and binary cross-entropy (BCE) as loss function and Intersection over Union (IoU) score as a performance metric was used. Binary cross-entropy is a loss function for binary classification tasks. The equation to calculate BCE can be seen in eq. 1. In the equation, G refers to the ground truth image, G_x refers to each pixel of that image, P refers to the predicted image and P_x refers to each pixel in the predicted image from the trained model. It is quite intuitive to interpret, loss score of 1 means predicted and ground truth does not match and if the score is 0 then output from the model precisely matches ground truth. IoU is a performance metric to measure the success of segmentation task than pixel wise accuracy in sparse image dataset which includes 80–90% background and very small percentage as positive labels or region of interest. IoU is independent of the true negative and therefore is more effective than naive measure such as accuracy. It measures the overlap between ground truth and predicted image. The equation to calculate IoU can be seen in eq. 2. The IoU score ranges between 0 and 1. A score of 1 indicates that the predicted image precisely matches with ground truth and a score of 0 indicates a complete mismatch. The sigmoid activation function was preferred over the softmax activation function because the model architecture has one neuron in the output layer to make it computationally more effective compared to the more generalized softmax function [30].

$$L_{BCE} = \sum_x -(G_x \log(P_x) + (1 - G_x) \log(1 - P_x)) \quad (1)$$

$$IoU = \frac{|G \cap P|}{|G \cup P|} \quad (2)$$

The images used for training were resized from 1525 \times 1128 pixels to 1024 \times 1024 pixels. The training of the model is largely supported by data augmentation to create new training samples from the existing ones. Data augmentation is a technique to increase the size of the dataset and prevent the model from overfitting. Albumentations is a fast and robust Python library for image data augmentation that features more than 70 different augmentations that include both pixel level and spatial level transformations [31]. Transformations such as horizontal flip, affine transforms, perspective transforms, brightness manipulations, contrast manipulations, image blurring and sharpening, Gaussian noise and random crop were applied as augmentation step to create a robust training dataset for U-net. The model was trained for 600 epochs with a batch size of 8 and a learning rate of 0.0001. The trained U-net model achieved an IoU score of 0.95 with a binary cross-entropy loss of 0.05 on the test set. The post-processing step involves the removal of artifacts and improves the grain boundary detection. This is performed automatically by morphological operations such as erosion and dilation.

All of the experiments in this paper were carried out using Python version 3.7 and other open-source libraries. For traditional machine learning, we used scikit-learn version 0.22.1 [32], and for deep learning Keras version 2.2.0 [33] with Tensorflow [34] backend. The deep learning model was trained on the system with a 256 GB NVIDIA Quadro RTX 8000 graphics card.

5. Results and discussion

To evaluate the performance of both approaches statistically, the output from the trained models was post-processed to obtain edge maps to measure morphologies such as area, equivalent circle diameter (ECD), aspect ratio, ferret maximum and circularity of each grain, and to obtain area-weighted and number-weighted grain size distribution curves based on equivalent circle diameter.

Fig. 10 shows the area-weighted and number-weighted grain size distribution of the samples NdFeB-C and NdFeB-D using a trained random forest classifier, U-net and manual approach. As described in Section 4.1, the results from the manual approach are considered as reference or ground truth data for performance comparison in this paper. Further, the area under the grain distribution curves obtained from different approaches is compared to assess the performance.

The area-weighted and number-weighted grain size distribution for samples NdFeB-C and NdFeB-D suggests that the distribution from the trained U-net model on both samples is in close range with the reference/manual values. The deviation of the area-weighted distribution

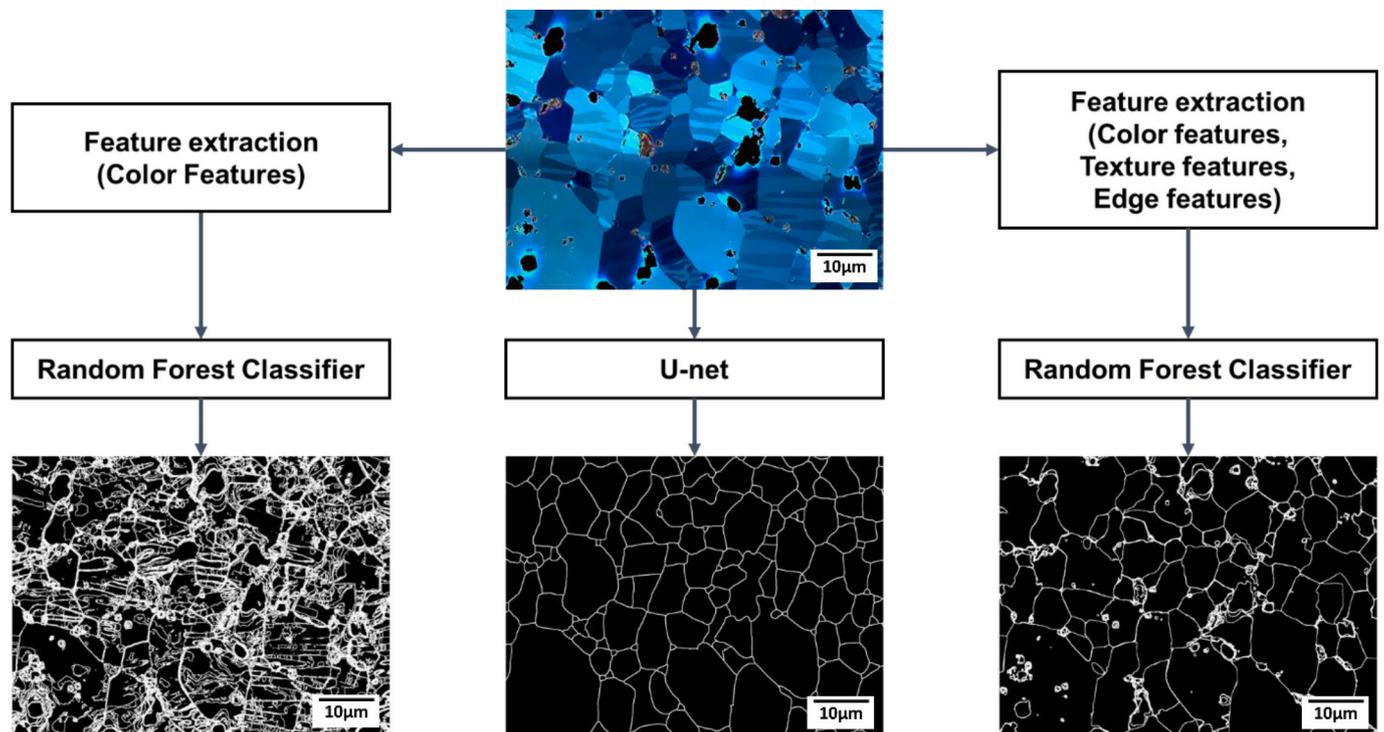


Fig. 9. Shows the primary difference between the workflow of the deep learning approach and the traditional machine learning approach for grain boundary detection from the Kerr image of the NdFeB sinter magnet sample. The output from the machine learning model trained on a lower-dimensional feature vector (color), a higher-dimensional feature vector (color, texture and edge) and a deep learning model is shown for visual comparison.

curve from the U-net model against the reference curve is 1.8% for NdFeB-C and 3% for NdFeB-D. A similar trend was observed when number-weighted distribution curves were compared. The deviation for NdFeB-C is 0.9% and for NdFeB-D it is 2.7%. The deviation for NdFeB-C predicted by the trained random forest classifier against the reference curve is 11% for the area-weighted distribution curve and 10% for the number-weighted distribution curve. However, for the NdFeB-D, sample deviation of the predicted area-weighted and number-weighted distribution curve by random forest classifier from the reference curve is 1.9% and 3% respectively. This is in close range of the reference values and output from the U-net model. One of the reasons for the better predictions by random forest classifier for NdFeB-D, when compared to NdFeB-C, is that some part of the NdFeB-D sample was included in the training dataset for both the random forest classifier and U-net model. Therefore, the random forest classifier had higher accuracy for the sample, which was part of the training dataset, whereas the U-net model performed well on both samples because features learned by the model here were highly complex and data augmentation helped the model to train on data with different noises, contrasts, etc.

As mentioned earlier, the results obtained from the manual approach involve the user tracing the grain boundaries from the Kerr images of the sample; the image is later processed using threshold-based image segmentation to calculate grain size distribution, which has been considered as ground truth or reference values. However, it is also important to evaluate the accuracy of the results obtained from the manual approach. For this, grain size distribution from the EBSD data was obtained for the NdFeB-A sample. EBSD has been a well-established and proven tool for the grain analysis of most materials, but it requires the sample to be prepared accordingly and it is also very time-consuming to generate results (including post-processing steps) [25].

Fig. 11 shows the area-weighted and frequency-weighted grain size distribution curve obtained using EBSD, manual, random forest classifier and U-net model. The difference in area under the area-weighted distribution curve between the manual and EBSD approach is 1.4%, and for the frequency-weighted distribution curve it is 0.5%. This

suggests that having a manual approach as a reference for comparing the performance of trained models is productive when compared to EBSD because the amount of time and effort involved in generating the results from the manual approach is less than using EBSD. Further, it has to be noted that it is a very challenging task to perform correlative microscopy involving EBSD and Kerr microscopy. This is primarily because it requires a high degree of technical skill to obtain the correlative image in magnet samples with the least amount of error.

The deviation of the area-weighted grain distribution curve against the reference curve obtained from the random forest classifier model is 1.9% and from the U-net model is 1.7%. Similarly, the deviation of the number-weighted grain distribution curve by random forest classifier is 2.5% and by U-net model is 2.1%. The performance of both models is nearly the same for this sample and is in close range to the reference values.

The traditional ML approach involves manual feature extraction; therefore the model requires the test images to have color or contrast in close range to that of the images used in the training dataset. This would mean that to have a robust model to work on different magnet samples, the training data should have images with different contrasts and colors. The microstructures of permanent magnets are very complex, and the presence of grains that are visible in Kerr images shows that feature extraction has to be very precise and of a higher dimension in order to have a robust feature-based ML model. To overcome this problem, the DL approach has been adopted to train a robust model because it does not require manual feature extraction; furthermore, using the data augmentation technique, it is possible to include data with different contrasts, colors and noises, etc. Fig. 12 shows the distribution curve for the NdFeB-B sample, which shows the results from the manual, EBSD and DL approach. The area-weighted distribution curve from the DL (U-net) model differs from the reference curve by 3.1% and the number-weighted distribution curve deviates from the reference curve by 4.3%. The deviation in the number-weighted curve is slightly higher than in the area-weighted curve and suggests that the U-net model detected more grains with a large equivalent diameter.

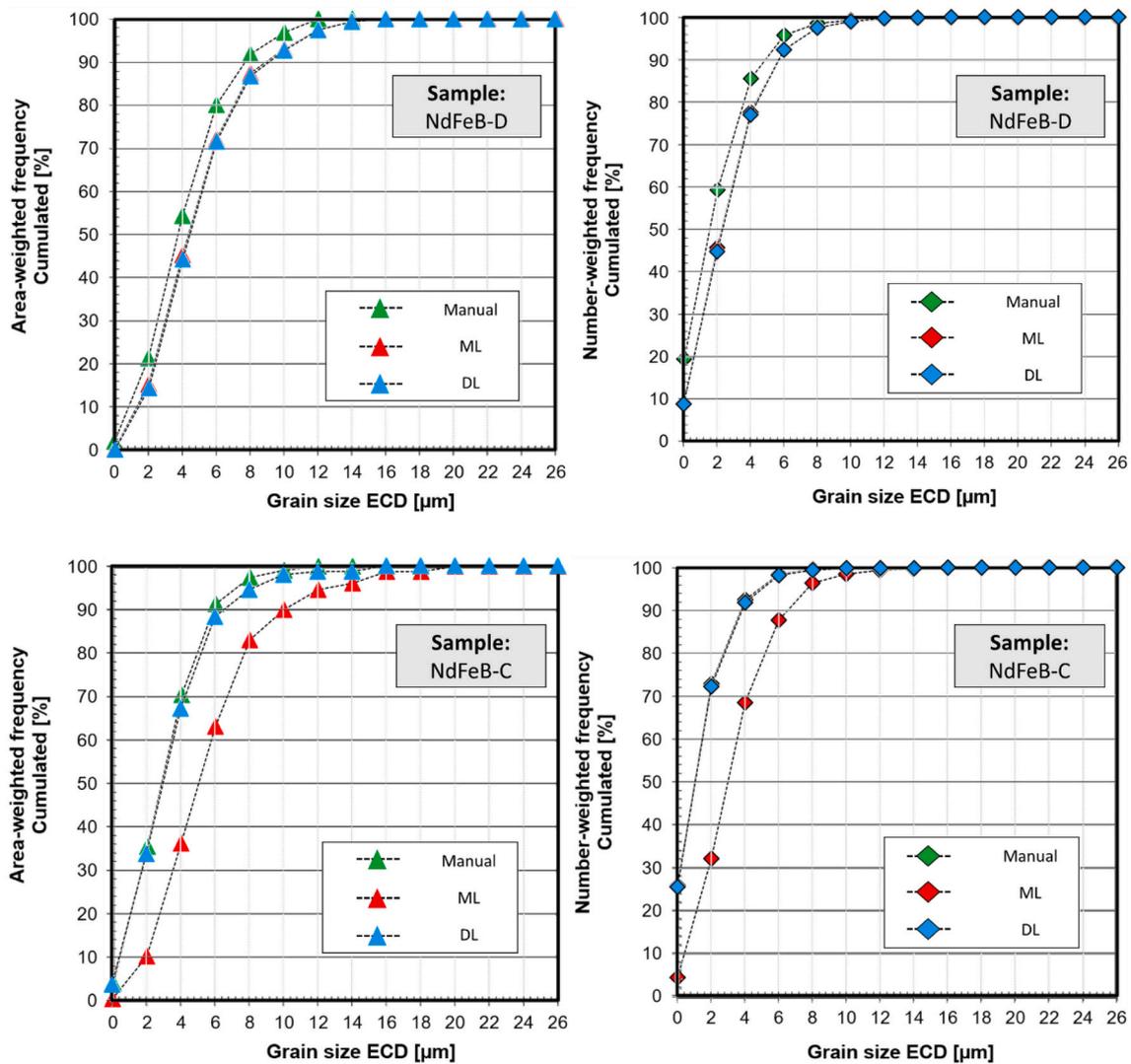


Fig. 10. Shows the area-weighted and number-weighted grain distribution curve for the NdFeB-C and NdFeB-D samples. The results from the trained machine learning and deep learning model are compared to the results from the manual (reference) approach.

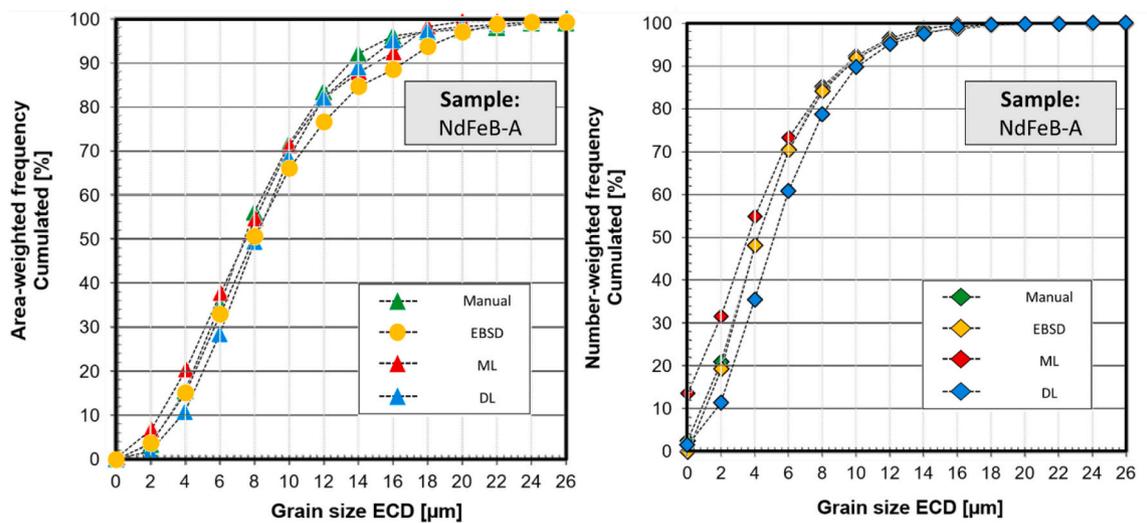


Fig. 11. Shows the area-weighted and number-weighted grain distribution curve for the NdFeB-A sample. The results from the trained machine learning model, deep learning model and EBSD are compared with the results from the manual (reference) approach.

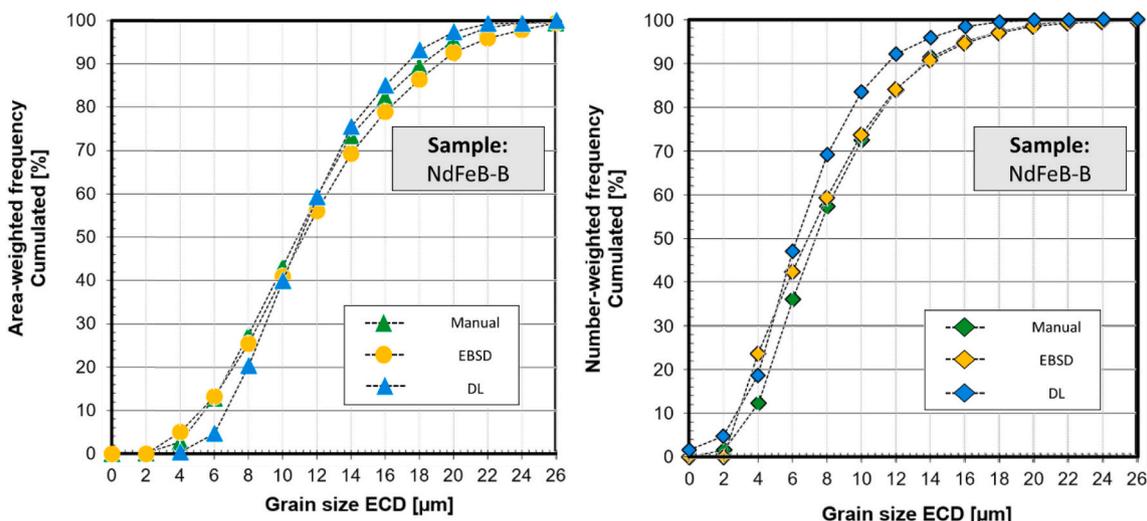


Fig. 12. Shows the area-weighted and number-weighted grain distribution curve for the NdFeB-B sample. The results from the trained deep learning model and EBSD are compared with the results from the manual (reference) approach.

However, some factors might affect the performance of the trained DL model. As mentioned in section 4.2 data augmentation techniques were included to improve the robustness of the model, yet it has some limitations concerning sample characteristics. The performance of the trained model would decrease or would fail to detect grain boundaries if

Kerr image resolution is too low, have strong preparation artifacts such as deep scratches passing through the grains and magnification of the acquired image is less than 500× for comparable grain sizes.

Additionally, the developed DL approach has some limitations when the magnet samples to be analyzed has a large number of grains with a

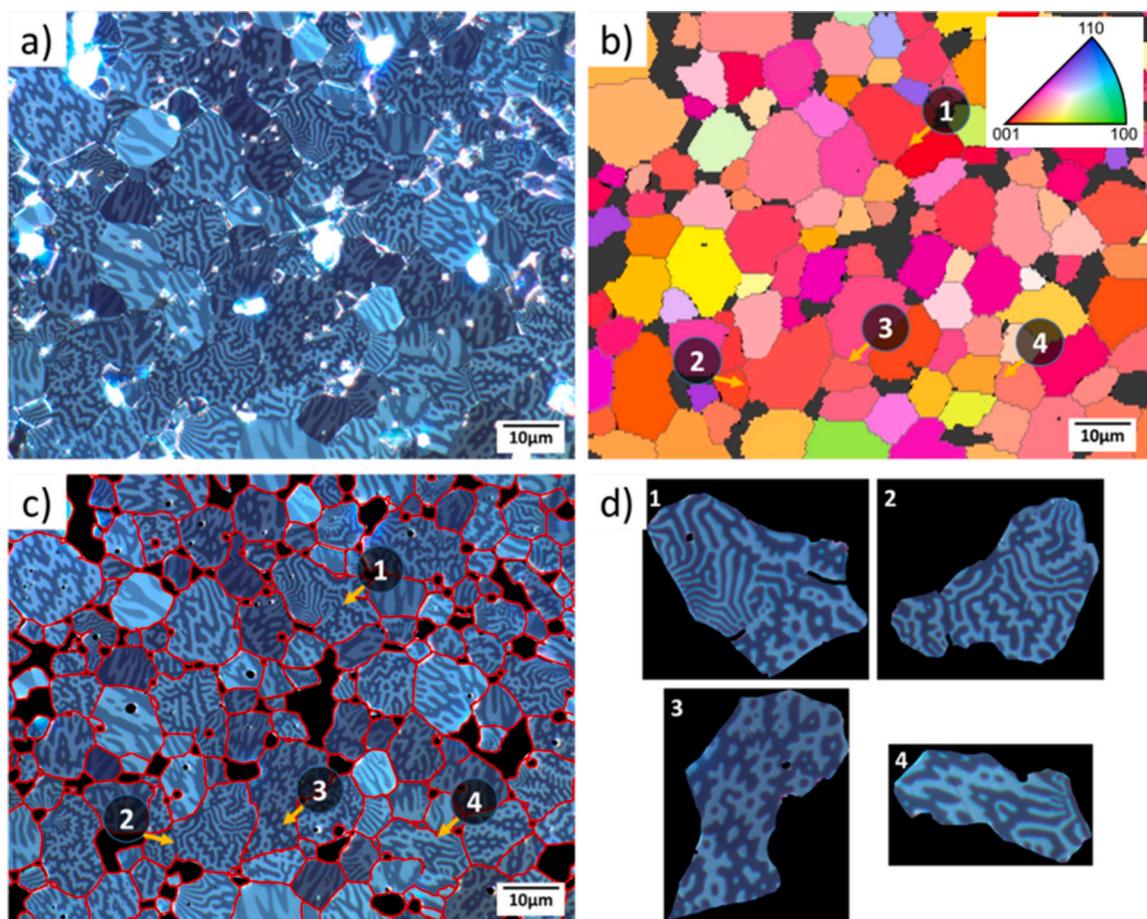


Fig. 13. Shows the visual comparison between the grain boundaries detected by different approaches in the $Nd_2Fe_{14}B_1$ sample with a high amount of grains with closure domain structures. (a) original image, (b) EBSD map (inverse pole figure) showing the grain boundaries, (c) grain boundaries detected by trained U-net in red color and, (d) individual images of extracted grains from trained U-net model where grain boundaries were not detected when compared to EBSD map.

closure magnetic domain structure. Fig. 1 shows the difference in the appearance of stripe and closure domains. For such samples, using the traditional approach completely fails and the manual approach would be very challenging. Fig. 13(c) shows the grain boundaries detected by the trained DL model on the sample with grains having closure domain structures. As a reference, an EBSD map is also generated for the same sample at the same position using correlative microscopy. The model fails to detect grain boundaries in some parts of the sample which can be seen in Fig. 13(b&c). For example, the EBSD map shows that there exists a grain boundary in grains shown in Fig. 13(d). One of the reasons for the failure of the trained model to detect some of the grain boundaries in the sample could be that the training dataset consists of images of samples that have more grains with stripe domains structures or a mixture of grains having stripe and closure domain structures. Further, it is a challenging task to manually label the images of a sample having a high number of grains with closure domain structures as they are prone to high false positives.

However, a deeper insight into the grain and its orientation can be performed using the EBSD tool that can measure the crystallographic orientation of grains in the samples and represent them as Bunge Euler angles Φ_1 (ϕ_1), Φ (Φ) and Φ_2 (ϕ_2) [25]. The difference between the crystallographic orientation of two grains/crystals with respect to each other is called misorientation. When the relative difference in ϕ_1 is high, the grain boundaries are visibly clear and this helps the trained model in better grain boundary detection. This effect gets stronger with the increase in the Φ value which is the reason why the trained model was able to detect grain boundaries in some grains within the sample in Fig. 13(a) and failed to detect grain boundaries in grains shown in Fig. 13(d). The relative difference in ϕ_1 value of the grains where the trained model failed to detect the grain boundary is low and also the relative misorientations ($\Delta\Phi$) of the grains in the range of 2° to 8° whereas other grains where model detected grain boundary have a high relative difference in ϕ_1 value despite having low Φ value. Moreover, it becomes challenging for the trained model to detect grain boundaries when relative misorientation is low. From the EBSD inverse pole figure map, which shows only Φ value it is evident that most of the grains in the sample in Fig. 13(a) have low Φ orientation. In-depth information on the EBSD tool for measurement on grain orientation and its application is not within the scope of this paper. Improving the current DL model to work effectively on samples with closure domain structures has been identified as part of future work.

6. Conclusion and future work

In this paper, a machine and deep learning approach for the detection of grain boundaries in NdFeB permanent magnets have been described. We have shown that both the traditional feature-based machine learning model and the state-of-the-art deep learning model were effective in terms of both accuracy and time by automatically detecting grain boundaries in large magnet samples.

The traditional machine learning-based model was equally effective as the deep learning model but lacks stability and robustness. The deep learning approach has proven to be more effective in terms of robustness and accuracy than the traditional machine learning approach. This is because the trained model can learn the meaningful features on its own by avoiding manual feature engineering. When compared to reference values, the overall deviation in the results from the trained U-net model is below 4%, and from the trained random forest classifier, it is between 3% to 11%. Further, the existing manual and EBSD approach consumes a relatively higher amount of time for grain analysis than the approaches described in this paper.

The U-net model discussed in the paper was able to automate the process of grain extraction and analysis from the Kerr microscopic images of the permanent magnet samples. Finally, we conclude that the tool described in this paper has the advantage of obtaining the grain size distribution from the light microscopic images with accuracy in close

range of the EBSD and the manual approach, and the time needed for analysis is greatly reduced. However, the trained model has limitations when the sample has more grains with closure domain structures and therefore this approach will perform with high accuracy and precision on the sample to be analyzed have grains that are primarily oriented with their magneto-crystalline anisotropy axis in the plane of the Kerr image axes.

Therefore, using ML techniques along with computer vision steps it is now possible to get grain size distribution directly from light microscopic images of magnetic materials, with higher accuracy and less processing time than EBSD and other manual approaches.

For future work, we would like to extend this approach to more permanent magnet alloys and also soft magnet alloys with stripe and closure domain structures. Also, we hope to extend this tool to include texture analysis, which would allow for comprehensive quality analysis to study different magnetic alloys.

Declaration of Competing Interest

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

Acknowledgements

The authors would like to thank the Magnets research group at Materials Research Institute Aalen, Germany, for providing her valuable feedback on the results. Additionally, the authors would like to thank the German Federal Ministry for Education and Research (BMBF) for funding this work under KLEVER project (Grant no: 13FH255PA6) as part of the FHProfUnt program.

References

- [1] M. Vilarinho Paula, *Functional materials: properties, processing and applications*, in: *Scanning Probe Microscopy: Characterization, Nanofabrication and Device Application of Functional Materials*, 2005, pp. 3–33.
- [2] T.T. Sasaki, T. Ohkubo, Y. Ue, H. Kubo, K. Hono, Effect of carbon on the coercivity and microstructure in fine-grained Nd-Fe-B sintered magnet, *Acta Mater.* 84 (2015) 506–514.
- [3] Wallace Matizanhuka, *The impact of magnetic materials in renewable energy related technologies in the 21st century industrial revolution: the case of South Africa*, *Adv. Mater. Sci. Eng.* 2018 (2018) 9.
- [4] J.F. Herbst, J.J. Croat, Neodymium-iron-boron permanent magnets, *J. Magn. Mater.* 100 (1991) 57–58.
- [5] James D. Widmer, Richard Martin, Mohammed Kimiabeigi, Electric vehicle traction motors without rare earth magnets, *Sustain. Mater. Technol.* 3 (2015) 7–13.
- [6] Toyota Global, Toyota Develops New Magnet for Electric Motors Aiming to [Online]. Available, <https://global.toyota/en/newsroom/corporate/21139684.html>, 20 February 2018 [Accessed 07 March 2021].
- [7] A. Hubert, R. Schäfer, *Magnetic Domains-The Analysis of Magnetic Microstructure*, Springer, 1998, pp. 11–97.
- [8] V. Soldatov, R. Schäfer, Selective sensitivity in Kerr microscopy, *Rev. Sci. Instrum.* 88 (2017), 073701.
- [9] V. Pusch, T. Bernthaler, F. Köppen, A. Wilde, A. Nagel, G. Schneider, Korrelative und quantitative Gefügeanalyse der Korn- und Domänenstruktur von Eisen-Neodym-Bor-Sintermagneten, *Sonderband der Praktischen Metallographie - Fortschritt in der Metallographie* 42 (2010) 171–176.
- [10] D. Goll, R. Loeffler, J. Herbst, R. Karimi, G. Schneider, High-throughput search for new permanent magnet materials, *J. Phys. Condens. Matter* 26 (6) (2014).
- [11] Lukas Exl, Johann Fischbacher, Alexander Kovacs, Harald Oezelt, Markus Gusenbauer, Kazuya Yokota, Tetsuya Shoji, Gino Hrkac, Thomas Schrefl, Magnetic microstructure machine learning analysis, *J. Phys. Mater.* 2 (2018), 014001.
- [12] R. Schäfer, Investigation of domains and dynamics of domain walls by the magneto-optical kerr-effect, in: *Handbook of Magnetism and Advanced Magnetic Materials*, John Wiley & Sons Ltd., US, 2007.
- [13] Jeffrey McCord, Progress in magnetic domain observation by advanced magneto-optical microscopy, *J. Phys. D. Appl. Phys.* 48 (2015), 333001.
- [14] A. Jansche, A.K. Choudhary, O. Badmos, H. Baumgartl, T. Bernthaler, R. Büttner, G. Schneider, Künstliche Intelligenz in der Materialmikroskopie, *Jahresmagazin Ingenieurwissenschaften / Werkstofftechnik* 2019 (2019) 14–17.
- [15] Saining Xie, Tu Zhuowen, Holistically-nested edge detection, in: *Proceedings of IEEE International Conference on Computer Vision*, 2015.

- [16] Olaf Ronneberger, Philipp Fischer, Thomas Brox, U-net: convolutional networks for biomedical image segmentation, *Medical Image Computing and Computer-Assisted Intervention (MICCAI)* 9351 (2015) 234–241.
- [17] Orkun Furat, Mingyan Wang, Matthias Neumann, Lukas Petrich, Matthias Weber, Cral E. Krill III, Volker Schmidt, Machine learning techniques for the segmentation of tomographic image data of functional materials, *Front. Mater.* 06 (2019) 123–139.
- [18] M. Neumann, R. Cabisco, M. Osenberg, H. Markötter, I. Manke, J.H. Finke, Characterization of the 3D microstructure of ibuprofen tablets by means of synchrotron tomography, *J. Microsc.* 274 (2019) 102–113.
- [19] Feng Jiang, Gu Qing, Huizhen Hao, Na Li, Grain segmentation of multi-angle petrographic thin section micrographic images, in: *International Conference on Image Processing (ICIP)*, China, 2017.
- [20] Wei Shen, Xinggang Wang, Yan Wang, Xiang Bai, Z. Zhang, DeepContour: A deep convolutional feature learned by positive-sharing loss for contour detection, in: *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, USA, 2015.
- [21] Gedas Bertasius, Jianbo Shi, Lorenzo Torresani, DeepEdge: a multi-scale bifurcated deep network for top-down contour detection, in: *IEEE Conference on Computer Vision and Pattern Recognition*, USA, 2015.
- [22] J. Joseph, C. Lim, Lawrence Zitnick, Piotr Dollar, Sketch tokens: a learned Mid-level representation for contour and object detection, in: *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, USA, 2013.
- [23] S. Xie, Z. Tu, Holistically-nested edge detection, in: *IEEE International Conference on Computer Vision (ICCV)*, Chile, 2015.
- [24] D. Martin, C. Fowlkes, D. Tal, J. Malik, A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics, in: *International Conference on Computer Vision*, USA, 2001.
- [25] Dominic Hohs, Tvrko Grubesa, David Schuller, Timo Bernthaler, Dagmar Goll, Gerhard Schneider, EBSD – a powerful tool for the analysis of magnetic materials, *Microsc. Microanal.* 23 (2017) 596–597.
- [26] David Schuller, Dominic Hohs, Ralf Loeffler, Timo Bernthaler, Dagmar Goll, Gerhard Schneider, Analysis of Soft Magnetic Materials by Electron Backscatter Diffraction as a Powerful Tool 8, *American Institute of Physics Advances*, 2018, 047612.
- [27] Dominic Hohs, David Schuller, Roland Stein, Timo Bernthaler, Dagmar Goll, Gerhard Schneider, Analyse maßgeblicher Gefügecharakteristika von Elektroband zur Minimierung magnetischer Verluste von Elektromotoren, in 51. Metallographietagung, Germany, 2017.
- [28] D. Hohs, D. Schuller, T. Grubesa, T. Schubert, R. Löffler, T. Bernthaler, D. Goll, G. Schneider, Quantitative Ermittlung von Gefügestrukturinformationen mittels korrelativer Mikroskopie und EBSD am weichmagnetischen Verbundwerkstoffen für elektrische Maschinen in 50. Metallographietagung, Germany, 2016.
- [29] E. Kleinberg, On the algorithmic implementation of stochastic discrimination, *IEEE Transact. PAMI* 22 (2000) 473–490.
- [30] Chigozie Nwankpa, Winifred Ijomah, Anthony Gachagan, Stephen Marshall, Activation Functions: Comparison of Trends in Practice and Research for Deep Learning, in: *International Conference on Computational Sciences and Technology (INCCST)*, Sindh Pakistan, 2018.
- [31] Buslaev Alexander, I. Iglovikov Vladimir, Khvedchenya Eugene, Parinov Alex, Druzhinin Mikhail, A. Kalinin Alexandr, Albuumentations: fast and flexible image augmentations, *Information* 11 (2020) 2078–2489.
- [32] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thiron, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, N. Brucher, M. Perrot, E. Duchesnay, Scikit-learn: machine learning in Python, *J. Mach. Learn. Res.* 12 (2011) 2825–2830.
- [33] F. Chollet, Keras, GitHub, 2015 [Online]. Available, <https://github.com/keras-team/keras>.
- [34] M. Abadi, A. Agarwal, P. Barham, E. Brevdo, Z. Chen, C. Citro, G. Corrado, A. Davis, J. Dean, M. Devin, TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems [Online]. Available, <https://www.tensorflow.org/>, 2015.