

Benchmarking deep learning for automated peak detection on GIWAXS data

Constantin Völter,^{a*} Vladimir Starostin,^b Dmitry Lapkin,^a Valentin Munteanu,^a Mikhail Romodin,^a Maik Hylinski,^a Alexander Gerlach,^a Alexander Hinderhofer^a and Frank Schreiber^a

Received 2 August 2024
Accepted 3 February 2025

^aInstitute of Applied Physics – University of Tübingen, Auf der Morgenstelle 10, 72076 Tübingen, Germany, and ^bCluster of Excellence ‘Machine learning – new perspectives for science’ University of Tübingen, Maria-von-Linden-Straße 6, 72076 Tübingen, Germany. *Correspondence e-mail: constantin.voelter@uni-tuebingen.de

Edited by H. Brand, Australian Synchrotron, ANSTO, Australia

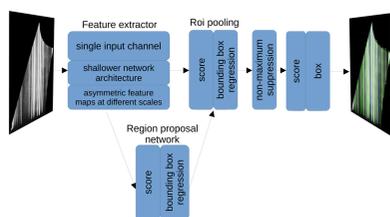
Keywords: deep learning; grazing-incidence wide-angle X-ray scattering; GIWAXS; peak detection; convolutional neural networks; Faster R-CNN.

Supporting information: this article has supporting information at journals.iucr.org/j

Recent advancements in X-ray sources and detectors have dramatically increased data generation, leading to a greater demand for automated data processing. This is particularly relevant for real-time grazing-incidence wide-angle X-ray scattering (GIWAXS) experiments which can produce hundreds of thousands of diffraction images in a single day at a synchrotron beamline. Deep learning (DL)-based peak-detection techniques are becoming prominent in this field, but rigorous benchmarking is essential to evaluate their reliability, identify potential problems, explore avenues for improvement and build confidence among researchers for seamless integration into their workflows. However, the systematic evaluation of these techniques has been hampered by the lack of annotated GIWAXS datasets, standardized metrics and baseline models. To address these challenges, we introduce a comprehensive framework comprising an annotated experimental dataset, physics-informed metrics adapted to the GIWAXS geometry and a competitive baseline – a classical, non-DL peak-detection algorithm optimized on our dataset. Furthermore, we apply our framework to benchmark a recent DL solution trained on simulated data and discover its superior performance compared with our baseline. This analysis not only highlights the effectiveness of DL methods for identifying diffraction peaks but also provides insights for further development of these solutions.

1. Introduction

The recent development of next-generation X-ray sources such as diffraction-limited synchrotrons and X-ray free-electron lasers has led to a drastic increase in the amount of data produced (Ludwig, 2019; Dong *et al.*, 2021; Qin & Bauer, 2010; Helliwell, 2019; Helliwell *et al.*, 2017). Furthermore, the latest generations of detectors enable experiments with high resolution in both the spatial and the temporal dimensions (Wang *et al.*, 2018; Heiss, 2019). The volume of data opens up new possibilities but makes manual data analysis infeasible. Large datasets demand the use of automated solutions for processing and analysis, while deep learning (DL) receives significant attention in the field (Hinderhofer *et al.*, 2023; Karniadakis *et al.*, 2021). DL stands out with its ability to comprehend intricate relationships, its capacity for generalization and its ability to handle vast amounts of data (Pithan *et al.*, 2023; Starostin *et al.*, 2022b; Greco *et al.*, 2022; Hinderhofer *et al.*, 2023; Guo *et al.*, 2016; Du *et al.*, 2016). A notable application of DL in this respect is the detection of Bragg peaks in X-ray scattering data, as demonstrated in the literature (Sullivan *et al.*, 2019; Liu *et al.*, 2022; Hadian-Jazi *et al.*, 2021; Yin *et al.*, 2022). The present paper evaluates the performance of the DL-based Bragg-peak-detection algorithm



proposed by Starostin *et al.* (2022a) and compares it with a conventional region-growing peak-finding algorithm. More specifically, these algorithms are applied to grazing-incidence wide-angle X-ray scattering (GIWAXS) data, which can be used to extract a wealth of structural information from thin films on the atomic scale (Feidenhans'l, 1989; Banerjee *et al.*, 2020; Robinson & Tweet, 1992). The ability to determine the arrangement and periodicity of crystal structures and their preferred orientations makes it indispensable in materials research and development (Steele *et al.*, 2023; Hu *et al.*, 2017; Brinkmann *et al.*, 2022). To extract this information, a key step is to accurately determine the positions, widths and intensities of Bragg peaks. However, expert knowledge is required to effectively fit the peaks due to the presence of diffuse background, experimental artefacts, scattering from the substrate and other environmental factors (Pauw, 2013). With hundreds of thousands of images captured per beam time, this laborious and time-consuming manual process becomes a bottleneck and presents a compelling case for implementing automated data analysis. The work by Starostin *et al.* (2022a) introduces such an automated peak detection using a two-stage object detection model. In the present study, we provide a test dataset, metrics and a baseline to evaluate the suggested approach. For our test dataset, we collected a diverse set of labelled GIWAXS data from perovskite films. This dataset encompasses a range of background signals, Debye–Scherrer rings with diverse widths and intensities, and Bragg peaks of varying shapes, intensities and proximities to one another. We find that traditional object detection metrics fail to capture critical aspects of peak-detection quality, which are essential for the subsequent analysis. As a result, we propose a physics-informed metric that prioritizes the radial over the azimuthal direction. For context, we implemented a baseline using a conventional region-growing algorithm (see *Appendix B* in the supporting information) commonly used for peak detection (Guo *et al.*, 2019). The introduced dataset, metrics and baseline were used for a comprehensive evaluation of the DL solution.

2. Dataset

2.1. Dataset

To facilitate a comparative analysis of peak detection in GIWAXS patterns, we curated a dataset collected during different steps of crystallization and annealing of various perovskite thin films prepared by spin coating as detailed by Kneschaurek *et al.* (2023). This includes various 2D and 3D perovskites with varying cations (caesium, methylammonium, formamidinium or their mixtures), anions (iodine, bromine) and spacer molecules [phenethylammonium, phenylenedimethylammonium, pentafluorophenylethylammonium or (1-adamantyl)methylammonium]. The selected patterns contain features corresponding to the final perovskite structures as well as different intermediate products (*e.g.* complexes with solvents) and precursors (lead iodide, lead bromide *etc.*). A glass slide covered with indium tin oxide (ITO) or fluorine-

doped tin oxide (FTO) and (optionally) mesoporous titanium oxide layers was used as a substrate. A full list of the selected compositions and structures is provided in the supporting information.

The dataset was measured at two experimental facilities: the P08 beamline at the Deutsches Elektronen-Synchrotron (DESY) and the ID10 beamline at the European Synchrotron Radiation Facility (ESRF). The X-ray energy varied between 18 and 22 keV, and the incidence angle varied between 0.1 and 0.5° (*i.e.* below and above the critical angle). Furthermore, our dataset features patterns that exhibit different resolutions, since they were measured using a PerkinElmer XRD 1621 detector with 2048 × 2048 pixels with a pixel size of 200 μm at the P08 beamline and a Pilatus 300K detector with 487 × 619 pixels with a pixel size of 172 μm at beamline ID10. For the latter, we merged several images of the same sample at different detector positions to fill the detector gaps.

2.2. Preprocessing

The raw data in detector coordinates contain Debye–Scherrer rings, in the form of arcs and arc segments, and Bragg peaks. This geometric configuration poses a challenge for the majority of computer vision algorithms, as many object detection techniques rely on rectangular bounding boxes and rectangular-shaped image filters. Consequently, we transform the GIWAXS images into a more computer-vision-compatible representation, utilizing a two-step image transformation. The GIWAXS image is first mapped from detector coordinates to sample-associated reciprocal space (Q_{\parallel} , Q_z) (Als-Nielsen & McMorrow, 2011), and then converted to polar coordinates ($|Q|$, ϕ) defined as

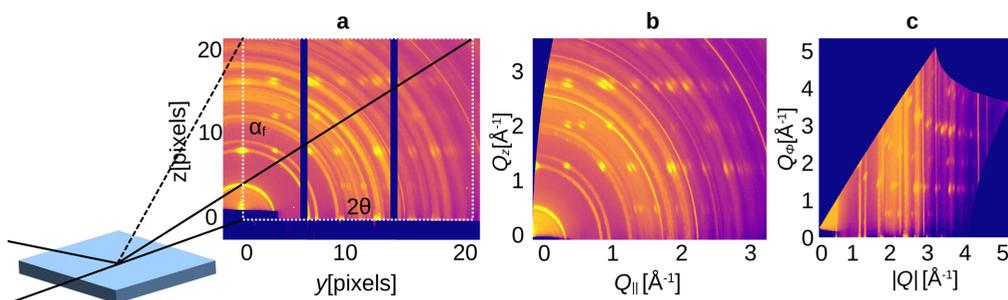
$$|Q| = (Q_z^2 + Q_{\parallel}^2)^{1/2}, \quad \phi = \arctan(Q_z/Q_{\parallel}), \quad (1)$$

where $|Q|$ is the radial coordinate and ϕ is the azimuthal angle counted from the sample horizon. For the region-growing approach, we discovered better performance with a conversion that does not distort the length of the peaks along the azimuthal direction. Instead, it focuses on the conservation of the azimuthal width of each peak, as shown in Fig. 1. We therefore chose the following conversion to ($|Q|$, Q_{ϕ}) defined as

$$|Q| = (Q_z^2 + Q_{\parallel}^2)^{1/2}, \quad Q_{\phi} = |Q| \arctan(Q_z/Q_{\parallel}), \quad (2)$$

where Q_{ϕ} is the distance along an azimuthal arc of radius $|Q|$ counted from the sample horizon. Note that Q_{ϕ} is not an actual scattering vector. The resulting images have a resolution of 1024 × 512. We provide an HDF5 file with the GIWAXS images in reciprocal space, along with a conversion script, on Zenodo (Völter *et al.*, 2024).

Depending on the specific measurement conditions and model of X-ray detector employed, the images exhibit unevenly distributed intensity levels, which can make it challenging to identify the peaks. To address this, we employed a contrast-limited adaptive histogram equalization (CLAHE) (Ketcham *et al.*, 1974), which distributes the intensities evenly


Figure 1

Geometry of GIWAXS experiments. Following the measurement, two acquired images are merged to remove the detector gaps and are subsequently converted from (a) detector coordinates to (b) reciprocal space coordinates and finally to (c) polar coordinates. For peak detection, the contrast is enhanced by CLAHE as described in Section 2.2. All shown images are already contrast-enhanced for visualization.

across the histogram. Note that the histogram equalization is applied exclusively for peak detection. Further analysis and fitting of the Bragg peaks is performed using the original data.

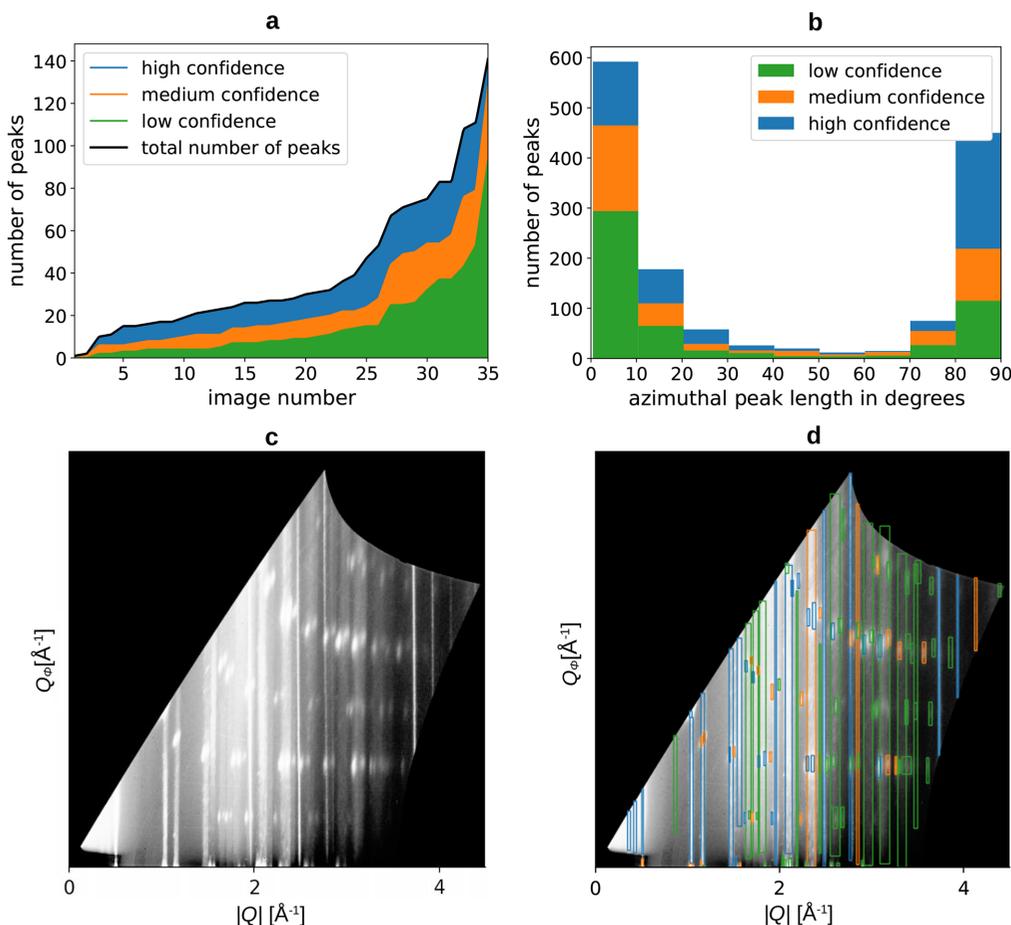
2.3. Annotation

To assess the precision of automated peak fitting, we manually annotated the radial and azimuthal positions of each Bragg peak in the described patterns with a bounding box. In

the radial dimension, we separately fitted each peak with a Gaussian function on top of a linear background:

$$I(|Q|) = I_0 \exp\left[-\frac{(|Q| - Q_{\text{fit}})^2}{2\sigma^2}\right] + B|Q| + C. \quad (3)$$

For the corresponding radial box width, we used the full width at half-maximum $w = 2(2 \ln 2)^{1/2} \sigma$ of the Gaussian function.


Figure 2

(a) Number of peaks per image in the evaluation dataset, sorted by the number of peaks. (b) Azimuthal length of peaks in the evaluation dataset. (c) and (d) Exemplary GIWAXS images from the dataset with labelled peaks. The colour is chosen according to the confidence level; the contrast is enhanced by CLAHE as described in Section 3.1.

In contrast, the peaks exhibit different shapes in the azimuthal direction, ranging from homogeneous segments of Debye–Scherrer rings to sharp and isolated Bragg peaks. This makes it impossible to use a single function for fitting. Consequently, we manually establish the boundaries for each peak in the azimuthal direction. We employ three confidence levels to characterize the prominence of the peaks. Bright peaks exhibiting a distinct Gaussian shape are assigned a high confidence rating, while peaks that are more challenging to discern receive a medium confidence level. The low-confidence category encompasses peaks that are scarcely visible and pose considerable difficulty in detection. They are barely above the background level or are covered by more intense neighbouring peaks. Fig. 2(a) shows both the peak count and the distribution of confidence levels of the 1448 peaks in the dataset. Fig. 2(b) shows the distribution of the azimuthal lengths of the peaks attributed to different confidence levels.

3. Automatic peak detection

3.1. Deep learning

The work by Starostin *et al.* (2022a) introduces a two-stage object detection model for peak detection and algorithms to further process the information obtained. We focus on the peak detection which employs a modified faster region-based convolutional neural network (Faster R-CNN) (Ren *et al.*, 2015) tailored to the GIWAXS geometry.

Ren *et al.* (2015) used convolutional neural networks (CNNs) to create feature maps from the given input image. These feature maps are abstract representations of the input image generated by a dot product operation with a convolutional kernel (LeCun *et al.*, 2015). We adapted the kernels of the CNN to be asymmetric, which reduces the image size mostly in the vertical direction. This elongated shape is well optimized for Debye–Scherrer rings and Bragg peaks, which

are typically broadened in the azimuthal direction due to the sample mosaicity.

For the second detection stage, the Faster R-CNN architecture uses feature maps of multiple scales. We modified this behaviour by including only the largest feature map. This prevents the network from confusing several distinct segments with one pronounced segment. A key part of the Faster R-CNN is the region-proposal network (RPN). It slides over the feature maps and creates proposals for potential objects. We pad the target boxes for the RPN such that the proposed regions contain more background area, which provides more context for the second detection stage. Furthermore, the RPN uses a reduced number of 64 channels. This shallower network architecture was found to be sufficient for this task. Additional customizations leverage the grey-scale nature of the images by using a single colour input channel. Since Bragg peaks are the only relevant class, the classifier component of the network is eliminated. These modifications lead to a substantial increase in processing speed while maintaining a high degree of accuracy. Fig. 3 shows the general structure of the Faster R-CNN and highlights the modifications of Starostin *et al.* (2022a).

3.2. Region-growing approach

To compare the performance of the DL approach against a standard solution, we developed a conventional algorithm based on a region-growing approach. Though tailored for the specific analysis of GIWAXS patterns, this algorithm serves as an exemplary demonstration of the differences between conventional and DL approaches. To provide an overview of the algorithm, Fig. 4 outlines the workflow.

Since this peak-detection method is primarily based on the difference in brightness between the peaks and the background, it is essential to smooth the background while keeping the shape of the original peaks [Fig. 4(a)]. In our experiments, we discovered that the Gaussian and box filters produce an

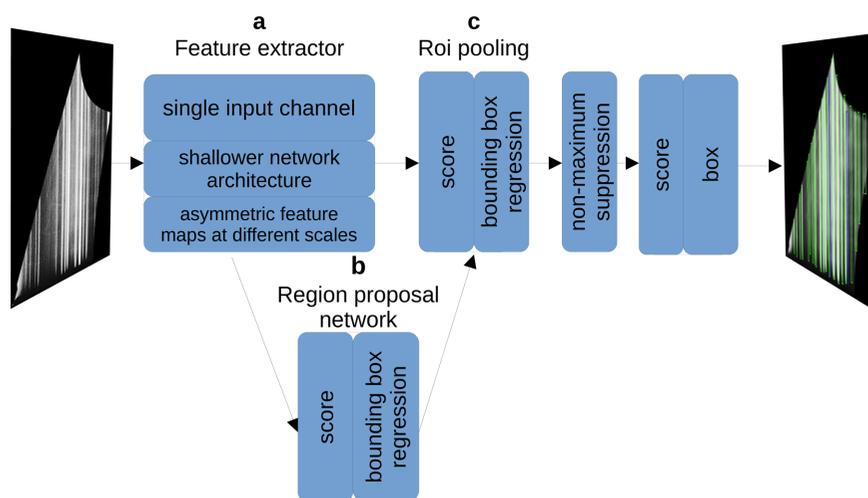


Figure 3

Overview of the modified Faster R-CNN structure. The feature extractor in (a) produces feature maps of different sizes and hands them on to the RPN in (b). This layer extracts regions of interest (RoIs) from the feature maps at different scales. The pooling layer in (c) aligns the proposed boxes from (b) and the largest feature map in (a). Non-maximum suppression eliminates overlapping boxes, resulting in the final predicted boxes together with their confidence scores.

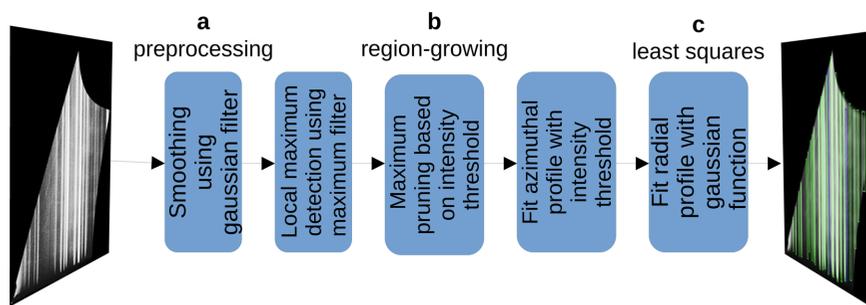


Figure 4 Peak-detection pipeline using a region-growing algorithm. (a) Image is smoothed using a 3×25 pixel Gaussian filter and a 3×3 pixel box filter. (b) The region-growing implementation detects peaks and fits their profile in the azimuthal direction. (c) A least-squares algorithm fits the radial profile using a Gaussian function with a linear background.

optimal blurred image for further processing. This denoising process significantly impacts the peak intensity and can noticeably shift the peak positions. Consequently, accurate peak positions must be determined through a peak fitting on the original, non-denoised image data.

As explained in Section 2.2, we used a different conversion to polar coordinates. This geometry is chosen because the distortion along the azimuthal direction presents a significant challenge for the region-growing algorithm (see Fig. 5). We find the geometry in Fig. 5(b) unsuitable for the region-growing algorithm due to the high number of false detections in the low- $|Q|$ range. We believe this is a result of the blurring

filters applied after the interpolation. Though noise of only a few pixels can be blurred, noisy pixels in the low- $|Q|$ region can expand in the polar coordinates and cannot be smoothed by the blurring. As a consequence, the region-growing algorithm detects them as maxima.

In the next step, the preprocessed image is utilized to detect intensity maxima in a two-way approach. Initially, local maxima are detected [Fig. 4(b)], and subsequently, the most prominent ones are selected among them [Fig. 4(b)]. We employ the Python implementation of Waithe (2023) which was initially proposed by Rueden *et al.* (2017).

To identify the local maxima, the algorithm employs a 3×3 pixel maximum filter, comparing the highest value within a 3×3 pixel neighbourhood with the corresponding values in the unfiltered image. Locations with identical values are identified as local maxima. Once the local maxima have been identified, the algorithm finds the global maxima across the entire image. This is achieved by growing the region around a maximum until an intensity threshold of 14 is met. This threshold corresponds to 5.5% of the absolute image brightness. The algorithm then combines the local maxima of a region into a single maximum. We determined the tunable parameters shown in Table 1 through systematic experimentation to achieve the highest recall values on the dataset described in Section 2.

To compare with the work of Starostin *et al.* (2022a), relying solely on the approximate peak position is not sufficient. In the radial direction, we employed the least-squares method to fit each peak with a Gaussian function on top of a linear background [Fig. 4(c)]. The boundary in the azimuthal

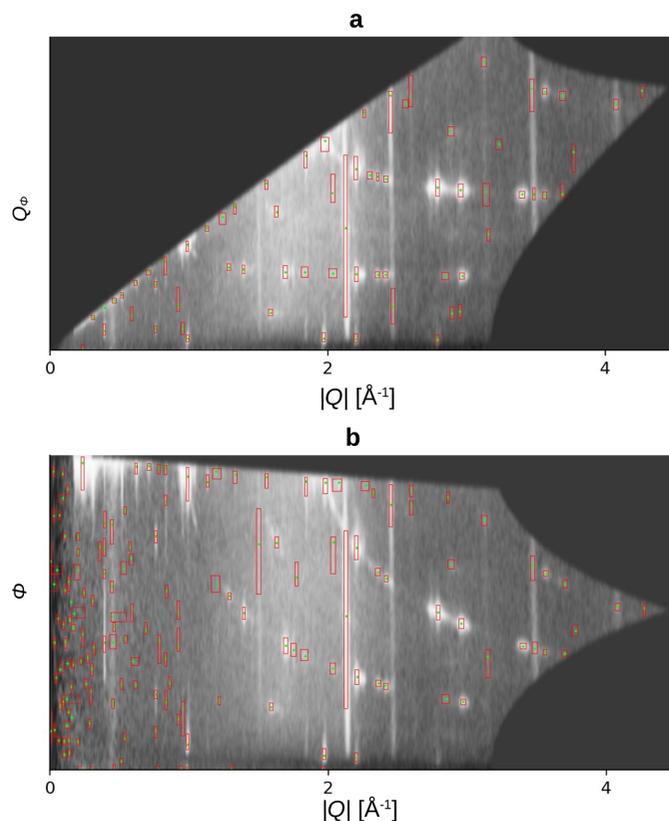


Figure 5 Detection results of the region-growing algorithm on the (a) undistorted and (b) distorted polar geometry. The red boxes and green points show a detected peak.

Table 1
Tunable parameters of region-growing approach.

Parameter	Value
Size of input image	1024 × 512 pixels
Filters used for image smoothing	Gaussian filter with kernel size 3×25 Box filter with kernel size 3×3
Kernel size of maximum filter	3×3 pixels
Threshold for region-growing algorithm	5.5% of image intensity
Threshold for fit in azimuthal direction	5.5% of image intensity
Fitting function in radial direction	Gaussian function

direction is determined by the region-growing algorithm using the intensity threshold of 14 [Fig. 4(b)].

4. Metrics

To evaluate the effectiveness of the two peak-detection methods, it is necessary to employ a metric designed specifically for Bragg peak detection. Hence, we will briefly examine metrics from related areas and tailor them to our specific use case. Hinderhofer *et al.* (2023) classify the task of Bragg peak detection as an object detection problem in the context of computer vision. This categorization is advantageous as it provides access to a range of well established metrics for evaluating and quantifying performance (Padilla *et al.*, 2020). One commonly used metric is the average precision, which assesses the accuracy by estimating the area under the precision–recall curve. The fundamental concept revolves around determining the intersection over union (IoU) criterion between the predicted box and the ground truth box:

$$\text{IoU} = \frac{\text{area of overlap}}{\text{area of union}}. \quad (4)$$

The identified bounding boxes are subsequently categorized as true positives, false positives or false negatives on the basis of the selected threshold for the IoU criterion (see Table 2).

The precision P and recall R are then determined by calculating the ratio of correct positive detections over all detections and all ground truths, as

$$P = \frac{\text{TP}}{\text{TP} + \text{FP}}, \quad (5)$$

$$R = \frac{\text{TP}}{\text{TP} + \text{FN}}, \quad (6)$$

where TP, FP and FN are the numbers of true positives, false positives and false negatives, respectively. The precision represents the algorithm’s ability to accurately identify relevant objects among the predicted instances. On the other hand, the recall quantifies the algorithm’s capacity to identify all of the given ground truth instances. Different confidence scores of the model result in different trade-offs between precision and recall points, which form a precision–recall curve $P(R)$. A high area under the curve indicates that many objects are recalled with high-quality intersections. Instead of integrating the whole area, a traditional approach is to interpolate the shape at 11 precision values (Schütze *et al.*, 2008). The result is the average precision

$$\text{AP} = 1/11 \sum_{R \in \{0.0, 0.1, \dots, 0.9, 1\}} P(R), \quad (7)$$

where $P(R)$ is the precision value P at the corresponding recall value R .

4.1. GIWAXS-specific metrics

While equation (7) is a commonly used metric for standard object detection tasks, we have identified the need for some adjustments to suit our specific peak-detection methods

Table 2

Classification of detection results based on the IoU criterion.

True positive (TP)	Detection successful: IoU threshold for intersection with ground truth box is met
False positive (FP)	Detected box has not met an IoU intersection threshold with a ground truth box
False negative (FN)	No box found meets the IoU threshold of the ground-truth box

better. The rationale stems from the features typically observed in GIWAXS patterns. The GIWAXS patterns contain Bragg peaks with a small, sharp Gaussian shape in the radial direction, whereas the azimuthal width is typically substantially larger. For determining the crystalline structure, the radial peak position is more important and requires a significantly higher level of precision in the radial direction compared with the azimuthal one. The IoU as a criterion of intersection is only partially suitable for this purpose since it treats both directions equally. Furthermore, the average precision is a single number to determine the quantity of intersections and only partially takes into account the quality of intersections. Therefore, we propose splitting the metric as explained in the following. The accurate determination of intensity requires a robust fit in both the radial and the azimuthal angles. Therefore, the average IoU ($\langle \text{IoU} \rangle$) is well suited for this specific purpose. To place additional emphasis on the radial direction, we propose to compute the IoU based only on the overlap in the radial direction (IoU_r). The average IoU in the radial direction ($\langle \text{IoU}_r \rangle$) reliably measures the quality of intersection among detected peaks. Given a specific IoU_r threshold, the average precision can be determined using the P_r and R_r values:

$$\text{AP}_r = 1/11 \sum_{R_r \in \{0.0, 0.1, \dots, 0.9, 1\}} P_r(R_r). \quad (8)$$

We adhere to the conventional 11-point interpolation because of its clear computational advantage in estimating the shape of the entire area under the curve.

Given the significance of the peak positions in the radial direction $|Q|$, the average distance between the detected $|Q_{\text{detected}}|$ and ground truth $|Q_{\text{truth}}|$ peak positions –

$$\langle \Delta |Q| \rangle = \langle ||Q_{\text{detected}}| - |Q_{\text{truth}}|| \rangle, \quad (9)$$

where the averaging is performed over all detected peaks – is of particular interest.

5. Results and discussion

We applied the peak-detection methods described in Section 3 with the composed dataset described in Section 2 and evaluated the results using the proposed metrics. For the classification of peaks as TPs, a minimum IoU_r value of 0.1 was used. Since the region-growing approach does not have a confidence score, we did not calculate the average precision but used the P_r of equation (5) based on a minimum IoU_r value of 0.1. The P_r value of the modified Faster R-CNN is calculated for a minimum confidence score of 0.1 and an IoU_r value of 0.1. The

results are summarized in Table 3. We observe that the DL approach outperforms the region-growing algorithm in almost all metrics.

Using the region-growing approach as a reference, the ‘Recall’ metric demonstrates that the Faster R-CNN approach detects more peaks in the high and medium confidence levels, whereas the region-growing approach has a slightly higher value for the low-confidence peaks. The detection of weaker peaks poses a challenge for both methods. Although the AP_r cannot be determined for the region-growing algorithm, the modified Faster R-CNN shows a promising value of 70% on experimental data. Note that the DL model is trained and fine-tuned on simulations, for which it reaches an AP_r value of 99% (Starostin *et al.*, 2022a). This gap could be reduced by including annotated, experimental data into the fine-tuning process. When contrasted with the region-growing approach, the Faster R-CNN shows a significantly higher P_r of 87% as opposed to 62%. Although both approaches have the same IoU_r in the radial direction, the average peak distance $\langle \Delta|Q| \rangle$ shows that DL detects peak centres more precisely. The IoU shows the same behaviour in the azimuthal direction: the DL approach demonstrates a superior fit.

Fig. 6 is an example from the evaluated dataset, visually confirming the results in Table 3. It demonstrates that the

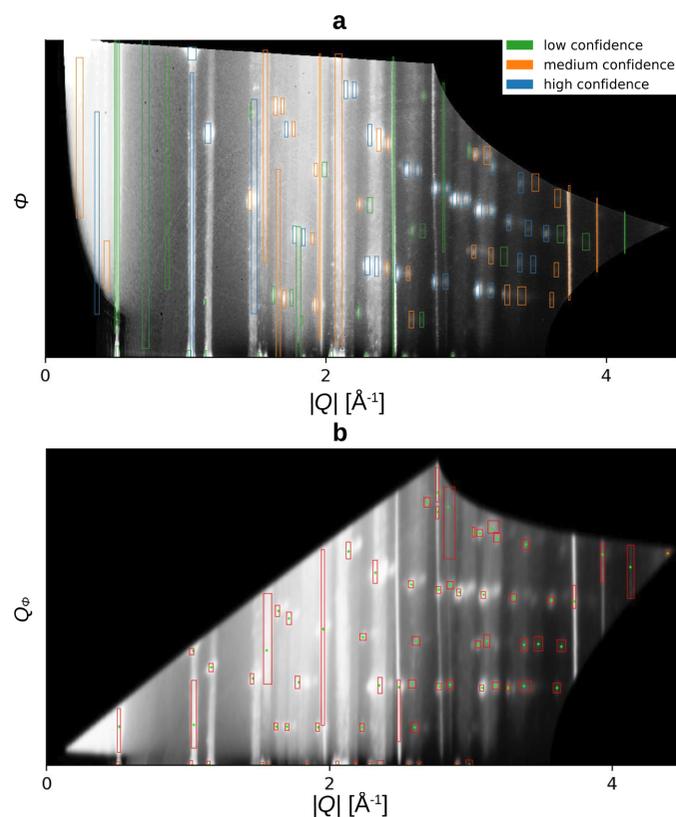


Figure 6 Peak-detection results from (a) the modified Faster R-CNN and (b) the region-growing approach. The colours in (a) are chosen according to the confidence score of the DL-model, not the confidence score of the ground truth labels. The different geometries are explained in Section 2.2, the blurring of (b) in Section 3.2. (b) Detected maxima are shown as green points. The red bounding boxes are the result of the fitting explained in Section 3.2.

Table 3 GIWAXS-specific metrics for the peak-detection results in the composed dataset.

Better results are marked in bold.

Metric	Modified Faster R-CNN			Region-growing approach		
	High	Medium	Low	High	Medium	Low
Confidence Recall (%)	95	83	55	88	73	60
AP_r (%)		70			–	
P_r (%)		87			62	
$\langle IoU \rangle$ (%)		49			35	
$\langle IoU_r \rangle$ (%)		64			64	
Percentile $\langle \Delta Q \rangle$ (10^{-3} \AA^{-1})	5	50	95	5	50	95
	0.47	6.24	17.76	0.83	6.26	24.25

region-growing approach has a lower recall for high- and medium-confidence peaks. The $\langle IoU \rangle$ is lower for the region-growing approach due to significant issues with azimuthally extended peaks, especially with Debye–Scherrer rings. In contrast, the radial intersection over union ($\langle IoU_r \rangle$) performs equally well for both algorithms. Evaluating $\langle \Delta|Q| \rangle$ visually is challenging without seeing the ground-truth boxes. All evaluated GIWAXS images, including the marked detection results, are available on Zenodo (Völter *et al.*, 2024).

5.1. Recall

We consider Bragg peak detection as effective if it achieves the highest possible recall value. As we can see in Table 3, both algorithms excel at detecting the brighter, high-confidence peaks, with recall values of more than 88%, but have problems detecting the less intense low-confidence peaks, showing results of 55 and 60%. We explain these results by examining how the algorithms extract relevant features.

In our specific case of detecting Bragg peaks, the region-growing algorithm uses the intensity of the peaks as a feature. The expert must determine a suitable intensity threshold to distinguish meaningful peaks from noise and background artefacts. Unfortunately, this feature is highly vulnerable to background scattering; the intensity as a characteristic alone may not be sufficient to achieve excellent results. This task requires more sophisticated features to enhance the performance further.

In contrast, a DL model possesses an intrinsic ability to detect features from a dataset. The DL model employed by Starostin *et al.* (2022a) is trained on a simulated dataset with varying peak intensities and captures these features by employing millions of training parameters. Nonetheless, the recall metrics in Table 3 indicate that the model struggles to detect less prominent peaks. Hence, enhancing the performance of feature detection could be realized through either a broader range of simulations or modifications to the model that enable the identification of peaks with low intensity.

Bragg peaks in GIWAXS data can manifest at arbitrary spatial positions, including border regions of the image. It is not uncommon for peaks to be located close to each other or even be situated on top of the Debye–Scherrer rings. Ensuring

a high recall in peak detection demands a robust algorithm capable of handling these characteristics.

Both of the discussed peak-detection methods demonstrate the capability to identify Bragg peaks at various positions. The approach to GIWAXS peak detection introduced by Starostin *et al.* (2022a) uses a diverse training dataset encompassing peaks with multiple spatial arrangements and therefore enables the identification of Bragg peaks across the entire image. Similarly, the region-growing approach detects local maxima across the entire image, highlighting its translational equivariance in peak detection.

We observe that Bragg peaks often appear very close to each other, which can result in missing out on individual peaks or mistakenly combining two into one. To achieve a high recall value, we require an algorithm that can accurately detect even these closely spaced peaks. However, the non-maximum suppression mechanism inherent in the Faster R-CNN architecture introduces challenges when dealing with these adjacent peaks. If a neighbouring peak holds lower confidence compared with another peak but significantly overlaps with it, the non-maximum suppression may erroneously suppress it (Bodla *et al.*, 2017). Furthermore, the non-maximum suppression mechanism is not equipped to handle instances where objects are nested, as exemplified by a Bragg peak atop a Debye–Scherrer ring. In contrast, the region-growing approach distinguishes peaks from one another on the basis of intensity thresholds. This approach performs effectively for prominent peaks, but it encounters challenges when dealing with peaks that do not meet the intensity difference requirement.

5.2. Peak refinement and box fitting

The $\langle \text{IoU} \rangle$ value in Table 3 gives the intersection between the labelled and detected peak positions. We deem the values of 49 and 35% not good enough for subsequent Bragg peak analysis, such as intensity determination. We believe these values can be partially attributed to the ambiguity of the peak position determination in the azimuthal direction. The azimuthal intensity distribution is not always peaked but may be more complex. Therefore it is not always possible to determine a reliable peak position. Consequently, the automated solutions encountered difficulties with it. While both models leave room for improvement, the Faster R-CNN, trained on a diverse range of peak shapes, demonstrates superior capability in the azimuthal position determination.

In contrast, $\langle \text{IoU}_r \rangle$ in the radial direction exhibits significantly better results with the value of 64%. The models can fit the radial shape accurately due to the Gaussian-like shape of the peaks and ring segments, which reduces ambiguity. The accurate determination of the radial peak position is crucial since it directly determines the uncertainty of lattice parameter estimation. The median shows similar results for both methods of 6.24×10^{-3} and $6.26 \times 10^{-3} \text{ \AA}^{-1}$. However, the modified Faster R-CNN outperforms the region-growing approach in the 5th and 95th percentiles, achieving 0.47×10^{-3} and $17.76 \times 10^{-3} \text{ \AA}^{-1}$ versus 0.83×10^{-3} and $24.25 \times$

10^{-3} \AA^{-1} . The values for both approaches are deemed acceptable for further processing such as structure identification.

5.3. Real time analysis

Modern X-ray sources allow measurement of X-ray scattering patterns at frequencies from a few hertz to megahertz (Li *et al.*, 2024; Decking *et al.*, 2020; Buffet *et al.*, 2012). Manually analysing hundreds of images per second is unrealistic. Consequently, a conventional approach would involve selecting a single image and analysing it, which could take anywhere from minutes to several hours depending on the complexity of the patterns. In contrast, the high processing rate in automated data analysis unlocks new types of experiments such as closed-loop experiments (Pithan *et al.*, 2023).

The modified Faster R-CNN model evaluated here benefits from the optimized software packages that are readily available, leveraging massively parallel algorithms, which results in an impressive number of 122 images processed per second (Starostin *et al.*, 2022a). This represents a significant acceleration when contrasted with the sequential region-growing approach, which handles approximately one image per second.

6. Conclusions

We have established a comprehensive benchmark for automated Bragg peak detection in the GIWAXS geometry. We have provided a labelled dataset, which spans a wide spectrum of practically relevant scenarios. We have proposed a new metric tailored to the specifics of the GIWAXS data that focus on physically relevant aspects of the detection performance. It is well adapted for Bragg peak detection, emphasizing the overlap in the radial direction instead of employing plain IoU. As a baseline, we developed a region-growing algorithm that was fine-tuned on this dataset. Our framework based on a recent DL method exhibits superior performance compared with this classical baseline fine-tuned on the test dataset. The findings indicate that the DL-based approach particularly excels at identifying the peak positions and boundaries. Noteworthy advantages of the DL method lie in accurately determining the azimuthal profile. Detecting low-intensity peaks is challenging due to the complexity of differentiating them from the background. Furthermore, identifying adjacent peaks poses difficulties, particularly in the context of non-maximum suppression used for the Faster R-CNN. Future work could aim to enhance the simulation or incorporate experimental data for training purposes. Furthermore, DL architectures that do not use non-maximum suppression could prove beneficial.

Acknowledgements

We acknowledge DESY (Hamburg, Germany), a member of the Helmholtz Association HGF, for the provision of experimental facilities. Data were partially collected at PETRA III,

and we would like to thank Florian Bertram for assistance in using beamline P08. We also thank the ESRF for providing synchrotron facilities and Oleg Konovalov from beamline ID10 for support during beam times. Open access funding enabled and organized by Projekt DEAL.

Funding information

We thank Bundesministerium für Bildung und Forschung [grant Nos. 05K19VTA (ERUM-PRO); 05K19VTC (ERUM-PRO); VIPR (ERUM-DATA)]. Further we thank the Cluster of Excellence – Machine Learning for Science, funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy (EXC No. 2064/1, project No. 390727645), for support and access to computer resources. We also thank the DFG (grant No. 460248799 to DAPHNE4NFDI).

References

- Als-Nielsen, J. & McMorrow, D. (2011). *Elements of modern X-ray physics*, 1st ed. Wiley.
- Banerjee, R., Kowarik, S. & Schreiber, F. (2020). *Advanced characterization of nanostructured materials*, World Scientific series in nanoscience and nanotechnology, Vol. 21, pp. 49–95. World Scientific.
- Bodla, N., Singh, B., Chellappa, R. & Davis, L. S. (2017). *Proceedings of the IEEE international conference on computer vision (ICCV)*, pp. 5562–5570. IEEE.
- Brinkmann, K. O., Becker, T., Zimmermann, F., Kreusel, C., Gahlmann, T., Theisen, M., Haeger, T., Olthof, S., Tückmantel, C., Günster, M., Maschwitz, T., Göbelsmann, F., Koch, C., Hertel, D., Caprioglio, P., Peña-Camargo, F., Perdigón-Toro, L., Al-Ashouri, A., Merten, L., Hinderhofer, A., Gomell, L., Zhang, S., Schreiber, F., Albrecht, S., Meerholz, K., Neher, D., Stolterfoht, M. & Riedl, T. (2022). *Nature*, **604**, 280–286.
- Buffet, A., Rothkirch, A., Döhrmann, R., Körstgens, V., Abul Kashem, M. M., Perlich, J., Herzog, G., Schwartzkopf, M., Gehrke, R., Müller-Buschbaum, P. & Roth, S. V. (2012). *J. Synchrotron Rad.* **19**, 647–653.
- Decking, W., Abeghyan, S., Abramian, P., Abramsky, A., Aguirre, A., Albrecht, C., Alou, P., Altarelli, M., Altmann, P., Amyan, K., Anashin, V., Apostolov, E., Appel, K., Auguste, D., Ayvazyan, V., Baark, S., Babies, F., Baboi, N., Bak, P., Balandin, V., Baldinger, R., Baranasic, B., Barbanotti, S., Belikov, O., Belokurov, V., Belova, L., Belyakov, V., Berry, S., Bertucci, M., Beutner, B., Block, A., Blöcher, M., Böckmann, T., Bohm, C., Böhnert, M., Bondar, V., Bondarchuk, E., Bonezzi, M., Borowiec, P., Bösch, C., Bösenberg, U., Bosotti, A., Böspflug, R., Bousonville, M., Boyd, E., Bozhko, Y., Brand, A., Branlard, J., Briechle, S., Brinker, F., Brinker, S., Brinkmann, R., Brockhauser, S., Brovko, O., Brück, H., Brüdgam, A., Butkowski, L., Büttner, T., Calero, J., Castro-Carballo, E., Cattalotto, G., Charrier, J., Chen, J., Cherepenko, A., Cheskidov, V., Chiadini, M., Chong, A., Choroba, S., Chorowski, M., Churanov, D., Cichalewski, W., Clausen, M., Clement, W., Cloué, C., Cobos, J. A., Coppola, N., Cunis, S., Czuba, K., Czwalińska, M., D'Almagne, B., Dammann, J., Danared, H., de Zubiurre Wagner, A., Delfs, A., Delfs, T., Dietrich, F., Dietrich, T., Dohlus, M., Dommach, M., Donat, A., Dong, X., Doynikov, N., Dressel, M., Duda, M., Duda, P., Eckoldt, H., Ehsan, W., Eidam, J., Eints, F., Engling, C., English, U., Ermakov, A., Escherich, K., Eschke, J., Saldin, E., Faesing, M., Fallou, A., Felber, M., Fenner, M., Fernandes, B., Fernández, J. M., Feuker, S., Filipakopoulos, K., Floettmann, K., Fogel, V., Fontaine, M., Francés, A., Martin, I. F., Freund, W., Freyermuth, T., Friedland, M., Fröhlich, L., Fusetti, M., Fydrych, J., Gallas, A., García, O., Garcia-Tabares, L., Geloni, G., Gerasimova, N., Gerth, C., Geßler, P., Gharibyan, V., Gloor, M., Głowinkowski, J., Goessel, A., Gołębiewski, Z., Golubeva, N., Grabowski, W., Graeff, W., Grebentsov, A., Grecki, M., Grevsmuehl, T., Gross, M., Grosse-Wortmann, U., Grünert, J., Grunewald, S., Grzegory, P., Feng, G., Guler, H., Gusev, G., Gutierrez, J. L., Hagge, L., Hamberg, M., Hanneken, R., Harms, E., Hartl, I., Hauberg, A., Hauf, S., Hauschildt, J., Hauser, J., Havlicek, J., Hedqvist, A., Heidbrook, N., Hellberg, F., Henning, D., Hensler, O., Hermann, T., Hidvégi, A., Hierholzer, M., Hintz, H., Hoffmann, F., Hoffmann, M., Hoffmann, M., Holler, Y., Hüning, M., Ignatenko, A., Iichen, M., Iluk, A., Iversen, J., Iversen, J., Izquierdo, M., Jachmann, L., Jardon, N., Jastrow, U., Jensch, K., Jensen, J., Jeżabek, M., Jidda, M., Jin, H., Johansson, N., Jonas, R., Kaabi, W., Kaefer, D., Kammering, R., Kapitza, H., Karabekyan, S., Karstensen, S., Kasprzak, K., Katalev, V., Keese, D., Keil, B., Kholopov, M., Killenberger, M., Kitaev, B., Klimchenko, Y., Klos, R., Knebel, L., Koch, A., Koepke, M., Köhler, S., Köhler, W., Kohlstrunk, N., Konopkova, Z., Konstantinov, A., Kook, W., Koprek, W., Körfer, M., Korth, O., Kosarev, A., Kosiński, K., Kostin, D., Kot, Y., Kotarba, A., Kozak, T., Kozak, V., Kramert, R., Krasnikov, M., Krasnov, A., Krause, B., Kravchuk, L., Krebs, O., Kretschmer, R., Kreutzkamp, J., Kröplin, O., Krzysik, K., Kube, G., Kuehn, H., Kujala, N., Kulikov, V., Kuzminych, V., La Civita, D., Lacroix, M., Lamb, T., Lancetov, A., Larsson, M., Le Pinvidic, D., Lederer, S., Lensch, T., Lenz, D., Leuschner, A., Levenhagen, F., Li, Y., Liebing, J., Lilje, L., Limberg, T., Lipka, D., List, B., Liu, J., Liu, S., Lorbeer, B., Lorkiewicz, J., Lu, H. H., Ludwig, F., Machau, K., Maciocha, W., Madec, C., Magueur, C., Maiano, C., Maksimova, I., Malcher, K., Maltezopoulos, T., Mamoshkina, E., Manschwetus, B., Marcellini, F., Marinkovic, G., Martinez, T., Martirosyan, H., Maschmann, W., Maslov, M., Matheisen, A., Mavric, U., Meißner, J., Meissner, K., Messerschmidt, M., Meyers, N., Michalski, G., Michelato, P., Mildner, N., Moe, M., Moglia, F., Mohr, C., Mohr, S., Möller, W., Mommerz, M., Monaco, L., Montiel, C., Moretti, M., Morozov, I., Morozov, P., Mross, D., Mueller, J., Müller, C., Müller, J., Müller, K., Munilla, J., Münnich, A., Muratov, V., Napoly, O., Näser, B., Nefedov, N., Neumann, R., Neumann, R., Ngada, N., Noelle, D., Obier, F., Okunev, I., Oliver, J. A., Omet, M., Oppelt, A., Ottmar, A., Oublaïd, M., Pagani, C., Paparella, R., Paramonov, V., Peitzmann, C., Penning, J., Perus, A., Peters, F., Petersen, B., Petrov, A., Petrov, I., Pfeiffer, S., Pflüger, J., Philipp, S., Pienaud, Y., Pierini, P., Pivovarov, S., Planas, M., Pławski, E., Pohl, M., Polinski, J., Popov, V., Prat, S., Prenting, J., Priebe, G., Pryschecki, H., Przygoda, K., Pyata, E., Racky, B., Rathjen, A., Ratuschni, W., Regnaud-Campderros, S., Rehlich, K., Reschke, D., Robson, C., Roeber, J., Roggli, M., Rothenburg, J., Rusiński, E., Rybaniec, R., Sahling, H., Salmani, M., Samoylova, L., Sanzone, D., Saretzki, F., Sawlanski, O., Schaffran, J., Schlarb, H., Schlösser, M., Schlott, V., Schmidt, C., Schmidt-Foehre, F., Schmitz, M., Schmökel, M., Schnautz, T., Schneidmiller, E., Scholz, M., Schöneburg, B., Schultze, J., Schulz, C., Schwarz, A., Sekutowicz, J., Sellmann, D., Semenov, E., Serkez, S., Sertore, D., Shehzad, N., Shemarykin, P., Shi, L., Sienkiewicz, M., Sikora, D., Sikorski, M., Silenzi, A., Simon, C., Singer, W., Singer, X., Sinn, H., Sinram, K., Skvorodnev, N., Smirnov, P., Sommer, T., Sorokin, A., Stadler, M., Steckel, M., Steffen, B., Steinhau-Kühl, N., Stephan, F., Stodulski, M., Stolper, M., Sulimov, A., Susen, R., Świerblewski, J., Sydło, C., Syresin, E., Sytchev, V., Szuba, J., Tesch, N., Thie, J., Thiebault, A., Tiedtke, K., Tischhauser, D., Tolkiehn, J., Tomin, S., Tonisch, F., Toral, F., Torbin, I., Trapp, A., Treyer, D., Trowitzsch, G., Trublet, T., Tschentscher, T., Ullrich, F., Vannoni, M., Varela, P., Varghese, G., Vashchenko, G., Vasic, M., Vazquez-Velez, C., Verguet, A., Vilcins-Czvitkovits, S., Villanueva, R., Visentin, B., Viti, M., Vogel, E., Volobuev, E., Wagner, R., Walker, N., Wamsat, T., Weddig, H., Weichert, G., Weise, H., Wendorf, R., Werner, M., Wichmann, R., Wiebers, C., Wiencek, M., Wilksen, T., Will, I., Winkelmann, L., Winkowski, M., Wittenburg, K., Witzig, A., Wlk, P., Wohlenberg, T., Wojciechowski,

- M., Wolff-Fabris, F., Wrochna, G., Wrona, K., Yakopov, M., Yang, B., Yang, F., Yurkov, M., Zagorodnov, I., Zalden, P., Zavadtsev, A., Zavadtsev, D., Zhirnov, A., Zhukov, A., Ziemann, V., Zolotov, A., Zolotukhina, N., Zummack, F. & Zybin, D. (2020). *Nat. Photon.* **14**, 391–397.
- Dong, H., Butler, K. T., Matras, D., Price, S. W. T., Odarchenko, Y., Khatri, R., Thompson, A., Middelkoop, V., Jacques, S. D. M., Beale, A. M. & Vamvakeros, A. (2021). *npj Comput. Mater.* **7**, 74.
- Du, X., Cai, Y., Wang, S. & Zhang, L. (2016). *Proceedings 2016: 31st youth academic annual conference of chinese association of automation (YAC)*, pp. 159–164. IEEE.
- Feidenhans'l, R. (1989). *Surf. Sci. Rep.* **10**, 105–188.
- Greco, A., Starostin, V., Edel, E., Munteanu, V., Rußegger, N., Dax, I., Shen, C., Bertram, F., Hinderhofer, A., Gerlach, A. & Schreiber, F. (2022). *J. Appl. Cryst.* **55**, 362–369.
- Guo, W., Zhang, Y., Hu, X., Zhang, T., Liang, M., Yang, X. & Yang, H. (2019). *IEEE Access*, **7**, 169715–169723.
- Guo, Y., Liu, Y., Oerlemans, A., Lao, S., Wu, S. & Lew, M. S. (2016). *Neurocomputing*, **187**, 27–48.
- Hadian-Jazi, M., Sadri, A., Barty, A., Yefanov, O., Galchenkova, M., Oberthuer, D., Komadina, D., Brehm, W., Kirkwood, H., Mills, G., de Wijn, R., Letrun, R., Kloos, M., Vakili, M., Gelisio, L., Darmanin, C., Mancuso, A. P., Chapman, H. N. & Abbey, B. (2021). *J. Appl. Cryst.* **54**, 1360–1378.
- Heiss, A. (2019). *Comput. Softw. Big Sci.* **3**, 15.
- Helliwell, J. R. (2019). *Struct. Dyn.* **6**, 054306.
- Helliwell, J. R., McMahon, B., Guss, J. M. & Kroon-Batenburg, L. M. J. (2017). *IUCrJ*, **4**, 714–722.
- Hinderhofer, A., Greco, A., Starostin, V., Munteanu, V., Pithan, L., Gerlach, A. & Schreiber, F. (2023). *J. Appl. Cryst.* **56**, 3–11.
- Hu, Q., Zhao, L., Wu, J., Gao, K., Luo, D., Jiang, Y., Zhang, Z., Zhu, C., Schaible, E., Hexemer, A., Wang, C., Liu, Y., Zhang, W., Grätzel, M., Liu, F., Russell, T. P., Zhu, R. & Gong, Q. (2017). *Nat. Commun.* **8**, 15688.
- Karniadakis, G. E., Kevrekidis, I. G., Lu, L., Perdikaris, P., Wang, S. & Yang, L. (2021). *Nat. Rev. Phys.* **3**, 422–440.
- Ketcham, D. J., Lowe, R. W. & Weber, J. W. (1974). *Image enhancement techniques for cockpit displays*. Technical Report. Defense Technical Information Center, Fort Belvoir, VA, USA.
- Kneschaurek, E., Hinderhofer, A., Hofferberth, B., Scheffczyk, N., Pithan, L., Zimmermann, P., Merten, L., Bertram, F. & Schreiber, F. (2023). *Rev. Sci. Instrum.* **94**, 063901.
- LeCun, Y., Bengio, Y. & Hinton, G. (2015). *Nature*, **521**, 436–444.
- Li, Q., Xing, R., Li, L., Yao, H., Wu, L. & Zhao, L. (2024). *Artif. Intell. Chem.* **2**, 100045.
- Liu, Z., Sharma, H., Park, J.-S., Kenesei, P., Miceli, A., Almer, J., Kettimuthu, R. & Foster, I. (2022). *IUCrJ*, **9**, 104–113.
- Ludwig, A. (2019). *npj Comput. Mater.* **5**, 70.
- Padilla, R., Netto, S. L. & Da Silva, E. A. B. (2020). *Proceedings of the international conference on systems, signals and image processing (IWSSIP)*, pp. 237–242. IEEE. <https://ieeexplore.ieee.org/document/9145130/>.
- Pauw, B. R. (2013). *J. Phys. Condens. Matter*, **25**, 383201.
- Pithan, L., Starostin, V., Mareček, D., Petersdorf, L., Völter, C., Munteanu, V., Jankowski, M., Konovalov, O., Gerlach, A., Hinderhofer, A., Murphy, B., Kowarik, S. & Schreiber, F. (2023). *J. Synchrotron Rad.* **30**, 1064–1075.
- Qin, J. & Bauer, M. A. (2010). *J. Phys. Conf. Ser.* **256**, 012007.
- Ren, S., He, K., Girshick, R. & Sun, J. (2015). *Proceedings of the 28th international conference on neural information processing systems (NIPS'15)*, Vol. 1, pp. 91–99. MIT Press.
- Robinson, I. K. & Tweet, D. J. (1992). *Rep. Prog. Phys.* **55**, 599–651.
- Rueden, C. T., Schindelin, J., Hiner, M. C., DeZonia, B. E., Walter, A. E., Arena, E. T. & Eliceiri, K. W. (2017). *BMC Bioinformatics*, **18**, 529.
- Schütze, H., Manning, C. & Raghavan, P. (2008). *Introduction to information retrieval*. Cambridge University Press.
- Starostin, V., Munteanu, V., Greco, A., Kneschaurek, E., Pleli, A., Bertram, F., Gerlach, A., Hinderhofer, A. & Schreiber, F. (2022a). *npj Comput. Mater.* **8**, 101.
- Starostin, V., Pithan, L., Greco, A., Munteanu, V., Gerlach, A., Hinderhofer, A. & Schreiber, F. (2022b). *Synchrotron Radiat. News*, **35**, 21–27.
- Steele, J. A., Solano, E., Hardy, D., Dayton, D., Ladd, D., White, K., Chen, P., Hou, J., Huang, H., Saha, R. A., Wang, L., Gao, F., Hofkens, J., Roefsaers, M. B. J., Chernyshov, D. & Toney, M. F. (2023). *Adv. Energy Mater.* **13**, 2300760.
- Sullivan, B., Archibald, R., Azadmanesh, J., Vandavasi, V. G., Langan, P. S., Coates, L., Lynch, V. & Langan, P. (2019). *J. Appl. Cryst.* **52**, 854–863.
- Völter, C., Starostin, V., Hinderhofer, A., Lapkin, D., Romodin, M., Kneschaurek, E. & Schreiber, F. (2024). *Gixd data and corresponding fits for machine learning*, <https://doi.org/10.5281/zenodo.11545913>.
- Waithe, D. (2023). *Maxima finding algorithm recreated from implementation in ImageJ/Fiji*, <https://github.com/dwaithe/MaximaFinder>.
- Wang, C., Steiner, U. & Sepe, A. (2018). *Small*, **14**, 1802291.
- Yin, J., Zhang, G., Cao, H., Dash, S., Chakoumakos, B. C. & Wang, F. (2022). *Accelerating science and engineering discoveries through integrated research infrastructure for experiment, big data, modeling and simulation*, edited by K. Doug, G. Al, S. Pophale, H. Liu & S. Parete-Koon, pp. 244–256. Springer Nature Switzerland.