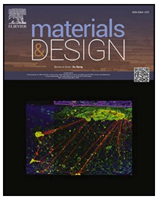
Materials & Design 263 (2026) 115669

Contents lists available at ScienceDirect

Materials & Design

journal homepage: www.elsevier.com/locate/matdes



Uncovering crystal structure evolution via nanobeam X-ray diffraction with a continuity-driven machine learning approach

Zhendong Wu[a,**], Tetsuya Tohei[a,**], Yusuke Hayashi[a,1], Shigeyoshi Usami[b], Masayuki Imanishi[b], Yusuke Mori[b], Junichi Takino[c], Kazushi Sumitani[d], Yasuhiko Imai[d], Shigeru Kimura[d], Akira Sakai[a,e,*]

[a] Graduate School of Engineering Science, The University of Osaka, 1-3 Machikaneyama, Toyonaka, Osaka 560-8531, Japan
[b] Graduate School of Engineering Science, The University of Osaka, 2-1 Yamadaoka, Suita 565-0871, Osaka, Japan
[c] Manufacturing Innovation Division, Panasonic Holdings Corporation, 1006, Kadoma, Kadoma City 571-8502, Osaka, Japan
[d] Japan Synchrotron Radiation Research Institute (JASRI), 1-1-1 Kouto, Sayo-cho, Sayo-gun, Hyogo 679-5198, Japan
[e] Spintronics Research Network Division, Institute for Open and Transdisciplinary Research Initiatives, The University of Osaka, 1-1 Yamadaoka, Suita 565-0871, Osaka, Japan

ARTICLE INFO

Keywords:
X-ray diffraction (XRD)
Unsupervised machine learning
Crystal
Microstructure
GaN

ABSTRACT

Nanobeam X-ray diffraction (nanoXRD) enables nanoscale mapping of crystal structures across wafer-scale crystals, offering unique insight into microstructural evolution during crystal growth. However, the resulting large and complex diffraction datasets make it challenging to quantitatively resolve local structural transitions and their connection to growth processes using conventional analysis. Here, we present a continuity-driven, unsupervised, and generalized analysis framework, referred to as the *neighborhood-based similarity metric*, which integrates spatial coordinates with nanoXRD data to reveal structural variations across growth sectors, interfaces, and defect-related regions without requiring prior knowledge or labels. By introducing Jaccard similarity scores to compare local neighborhoods in spatial and diffraction domains, the method quantitatively detects discontinuities where structural evolution disrupts the local continuity of diffraction patterns. Our unsupervised approach, validated with synthetic data and nanoXRD measurements of bulk GaN crystals, successfully identified both known defects and previously hidden structural discontinuities. The results provide new insights into the relationship between growth conditions, local strain evolution, and defect formation, establishing a robust and interpretable approach for linking processing and structural characteristics in complex crystalline materials.

1. Introduction

Nowadays, the functions of electronic devices increasingly rely on nanostructures [1,2], which motivates the need to map and localize crystal structural deformations with high spatial resolution and measurement efficiency, the latter being particularly important when studying bulk crystal substrates. Because real-time in situ monitoring of structural evolution during crystal growth is often impractical under extreme conditions such as high temperature, high pressure, or chemically reactive environments, such growth-induced structural evolution is typically investigated a *posteriori*. Spatially resolved structural variations are analyzed in cross-sectional, wafer-scale samples after growth to infer the evolution of the crystal structure during fabrication. From this perspective, the development of nanobeam X-ray diffraction (nanoXRD) and scanning X-ray diffraction microscopy (SXDM) has enabled non-destructive, quantitative characterization of crystal structures, with measurement areas easily adjustable to meet experimental demands [3–20]. However, the analysis of advanced nanoXRD datasets faces challenges arising from both the enormous number of diffraction patterns and the complex yet subtle structural variations in crystalline materials. The rapidly increasing rate of data acquisition and the growing complexity of the data can outpace the accuracy and robustness

---

* Corresponding author at: Graduate School of Engineering Science, The University of Osaka, 1-3 Machikaneyama, Toyonaka, Osaka 560-8531, Japan.
** Corresponding author.
E-mail addresses: u338462d@ecs.osaka-u.ac.jp (Z. Wu), tohei@ee.es.osaka-u.ac.jp (T. Tohei), sakai@ee.es.osaka-u.ac.jp (A. Sakai).
[1] Present address: National Institute for Materials Science (NIMS), 1-2-1 Sengen, Tsukuba, Ibaraki 305-0047, Japan.

https://doi.org/10.1016/j.matdes.2026.115669
Received 10 December 2025; Received in revised form 24 January 2026; Accepted 13 February 2026

of traditional analysis workflows based on conventional peak fitting [33]. While careful fitting with region-specific models and constraints can be effective at a local scale, such approaches do not readily scale to wafer-level discovery across structurally heterogeneous regions.

Machine learning (ML) methods offer a promising route to interpret such massive diffraction datasets by capturing correlations and anomalies that may correspond to hidden structural features [20–32]. Although most ML–XRD studies have focused on supervised learning, our previous work [33,34] demonstrated that unsupervised learning can also effectively reveal domain structures in bulk GaN crystals, classifying regions with distinct crystallographic characteristics. In wafer-scale nanoXRD/SXDM mapping, the primary objective is often the discovery and localization of microstructural transitions in the absence of reliable ground-truth labels and without assuming a universally valid parametric model for diffraction features. This objective motivates the unsupervised, continuity-driven formulation adopted here.

The success of unsupervised ML methods relies on the intrinsic relationship between the data structure and the crystal structure, whereby distinct crystal configurations yield characteristic variations in XRD signal features such as peak shape, position, width, and intensity. When pronounced differences exist between datasets, unsupervised methods, such as Uniform Manifold Approximation and Projection (UMAP) [35] and t-Distributed Stochastic Neighbor Embedding (t-SNE) [36], are effective in visualizing and identifying clusters in data, as well as interpreting distinct structural domains without prior labels.

Building on our previous findings on structural domains identified through clustering [34], a remaining challenge is to elucidate how crystal structures evolve across these domains during growth. In addition, we seek a generalized analysis strategy applicable to thousands of measurement points and different samples, particularly in situations where a single parametric peak-fitting model does not remain stable owing to variations in diffraction profile features. To address those challenges, the present study focuses on structural transitions that occur at interfaces and boundaries formed under different growth conditions. We introduce a continuity-aware, unsupervised analysis framework—the *neighborhood-based similarity metric*—which captures gradual changes in local structural continuity by integrating spatial and diffraction information. Applying this approach to multiple bulk GaN crystals enables quantitative visualization of structural continuity and the identification of interfacial regions that are not clearly resolved by conventional clustering or optical characterization. This framework not only complements domain-based analyses but also provides deeper insight into the mechanisms of interfacial formation, strain accommodation, and microstructural evolution during crystal growth, thereby linking various growth processing conditions to the resulting structural characteristics.

2. Methods

2.1. Preparation of raw data

This study investigated several bulk GaN crystal samples, including those prepared by the Na-flux, oxide vapor phase epitaxy (OVPE), and hydride vapor phase epitaxy (HVPE) methods. We conducted position-dependent nanoXRD experiments for those samples at the hard X-ray undulator beamline BL13XU in SPring-8. (Fig. 1) The photon energy was 8 keV with an energy resolution of 10^{-4}. The beam size was around 500 nm (hor.) x 390 nm (ver.). Here, hor. and ver. refer to the spatial horizontal and vertical directions, respectively. In Fig. 1(b), the spatial horizontal plane corresponds to the yz-plane, while the vertical direction is defined along the x-axis. For each sampling point, e.g., point *i* at (x_i, y_i), symmetric $(2\bar{2}00)$ and asymmetric $(2\bar{2}02)$ planes of GaN were chosen as diffraction planes. The diffracted X-rays from GaN crystals were detected by a two-dimensional (2D) photon-counting detector HyPix-3000 (Rigaku, Japan) and three-dimensional (3D) ω–2θ–φ data were collected by scanning incident angle ω, i.e., the ω-scan. Representative raw data of diffraction patterns are listed in *Supplementary Note 1*.

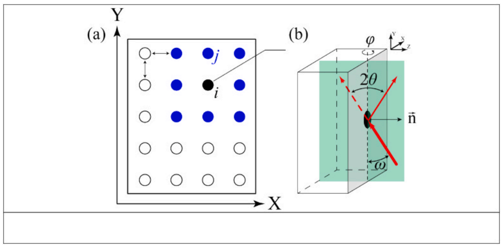


Fig. 1. Schematic of nanoXRD experiments. (a) Arrangement of measurement points on the sample surface. (b) Diffraction geometry, including the incident angle ω, diffraction angle 2θ, and the azimuthal angle φ. Under ideal diffraction conditions, the incident beam, the diffracted beam, and the normal of the diffracting crystal plane lie within the same plane, indicated by the green cross-section in (b).

2.2. Algorithm for analysis

The *neighborhood-based similarity metric* is based on the concept that, in a continuously and gradually varying crystal, neighboring measurement points (points *i* and *j* in Fig. 1(a)) share similar structures and therefore similar diffraction patterns. In other words, if two measurement points are spatial neighbors, the data obtained from them are expected to be neighbors in the data space. Conversely, when the crystal structure at neighboring points is deformed, i.e., when the structure undergoes an abrupt change, the corresponding data no longer remain neighbors in the data space.

In the *neighborhood-based similarity metric*, the position of point *i* is denoted as $S_i = (x_i, y_i)$, and the data obtained at this point is denoted as D_i. Furthermore, in contrast to the previous research, which relies solely on raw XRD intensities for clustering [33,34], the current method integrates S_i into the XRD features to compose the raw data. To further balance the effects of locations and data features in the calculation, the raw data is further composed of weighted XRD features of the form $D_i = (x_i,\ y_i,\ amp \bullet nor\_XRD_i)$, where the *amp* is a scalar weighting factor applied to the normalized XRD signal (nor_XRD_i) at point *i*. In this framework, the amp controls how strongly spatial relationships in real space constrain neighborhood construction relative to diffraction-pattern similarity. By adjusting amp, the method effectively tunes the extent to which structural discontinuities are emphasized in the results: larger values favor diffraction-driven similarity and enhance the contrast of weak deviations, whereas smaller values enforce stronger spatial coherence and suppress weak or non-local discontinuities.

During the calculation of the neighborhood-similarity metric, the set of neighboring points of *i* in the spatial space *S* is denoted as N_i^S, while the set of neighbors in the data space *D* of *i* is denoted as N_i^D. The *neighborhood-based similarity metric* consists of comparing the similarity between N_i^S and N_i^D, i.e., how much these two sets of neighbors overlap. In this study, the Jaccard similarity score *J* is utilized to quantitatively describe the similarity between N_i^S and N_i^D. A high score indicates that the spatial and data neighborhoods match well, suggesting strong local continuity near point i, as expected in regions where the crystal structure evolves smoothly and nearby points on the sample surface exhibit the most similar diffraction patterns. In contrast, a low score indicates a mismatch, suggesting a significant structural difference between nearby points. The detailed calculation procedure for the *neighborhood-based similarity metric* is presented in **ALGORITHM 1**, with an illustration in *Supplementary Note 2*.

ALGORITHM 1. The neighborhood-based similarity metric

```
procedure LOCALSIMILARITY(S, D, δx, δy)
```

(*continued on next page*)

(*continued*)

ALGORITHM 1. The neighborhood-based similarity metric

```
▷ Step 1: Local neighborhood comparison
2:  Input:
3:      $S \in \mathbb{R}^{N\times 2}$
4:      $D \in \mathbb{R}^{N\times(2+M)}$
5:      $\delta x, \delta y \in \mathbb{R}^{+}$
6:  Output:
7:      $J \in \mathbb{R}^{N}$
        ▷ Jaccard similarity for each point
8:      $N \leftarrow$ number of points in S, $J \leftarrow \varnothing$
9:      Dist ← DistanceMatrix(D)
            ▷ Pairwise distances matrix $(N \times N)$ in D
10:     for $i = 1$ to N do
11:         $(x_i, y_i) \leftarrow S[i]$
12:         $N_i^S \leftarrow$ FindNeighbors$(S, (x_i, y_i), \delta x, \delta y)$
            ▷ Neighbors for each point in spatial layout
13:         $d_{max} \leftarrow \max\{ Dist[i, j] \mid j \in N_i^S \}$
14:         $N_i^D \leftarrow \{ j \mid Dist[i, j] \le d_{max} \}$
15:         $J[i] \leftarrow | N_i^S \cap N_i^D | / | N_i^S \cup N_i^D |$
16:     end for
17:     return J
18: end procedure
19: function DistanceMatrix(P)
20:     $N \leftarrow$ number of points in P, $D \leftarrow$ matrix(N, N)
21:     for $i = 1$ to N do
22:         for $j = 1$ to N do
23:             $D[i, j] \leftarrow ||P[i] - P[j]||_2$
                ▷ Euclidean distance
24:         end for
25:     end for
26:     return D
27: end function
28: function FindNeighbors $(S, (x_i, y_i), \delta x, \delta y)$
29:     $N \leftarrow$ number of points in S, $N \leftarrow \varnothing$
30:     for $j = 1$ to N do
31:         $(x_j, y_j) \leftarrow S[j]$
32:         if $| x_i - x_j | \le \delta x$ and $|y_i - y_j| \le \delta y$ then
33:             $N \leftarrow N \cup \{j\}$
                ▷ Find points within a rectangular neighborhood
34:         end if
35:     end for
36:     return N
37: end function
```

As described in line 13 of the algorithm, the determination of N_i^D relies on a point-wise calculation of the neighborhood size in joint data space *D*. Specifically, the effective neighborhood extent around point *i* is defined by the Euclidean distances between point *i* and its spatial neighbors N_i^S, evaluated in the joint data space *D*. Since the *neighborhood-based similarity metric* quantifies the overlap between neighborhoods defined in spatial space *S* and in a joint data space *D*, it does not involve parametric peak fitting and does not require prior labels or model initialization. The calculation instead relies on the relative weighting parameter *amp* between spatial and diffraction features, rather than on peak-shape models and constraint tuning. Moreover, the proposed metric does not involve model training or loss minimization. Consequently, issues commonly associated with optimization-based machine-learning methods, such as overfitting or convergence to local minima, are not relevant in the same sense, and the calculation is deterministic and fully reproducible for fixed inputs and analysis settings. Finally, by avoiding the direct use of Euclidean distances in the full data space, the proposed metric reduces the influence of outliers arising from isolated points located far from the surrounding data distribution as well as regional biases within the data space.

Importantly, when constructing the joint data space *D*, both diffraction and spatial information contribute to the definition of neighborhoods. If diffraction patterns alone are used to define neighbors in *D*, the algorithm can identify globally similar patterns, but it loses the notion of locality that is essential for detecting interfaces as localized deviations from continuity. In particular, without spatial regularization provided by the (*X, Y*) component in *D*, diffraction patterns from non-neighboring regions, for example, distant areas belonging to a different growth sector but exhibiting similar diffraction patterns, may be selected as “neighbors” in the data space. This loss of spatial regularization introduces non-local similarities into the neighborhood comparison, thereby reducing the reliability of discontinuity localization, because the resulting score is no longer constrained to physically adjacent points within the sample.

3. Results and discussion

3.1. Simulation-based demonstration of methodology

To demonstrate that the *neighborhood-based similarity metric* enables the identification of structural variations, its performance is presented using a synthetic dataset, as shown in Fig. 2. A synthetic 2D measurement area of 30 × 40 μm^2 was defined, and the measurement point interval was set to 1 μm. We simulated the 2 [GRKSLT]-intensity spectrum from the $Al_xGa_{1\text{-}x}N$ $2\bar{2}00$ diffraction with $a_{GaN} = 3.189$ Å, $c_{GaN} = 5.186$ Å, $a_{AlN} = 3.112$ Å, $c_{AlN} = 4.982$ Å, and the energy of the incident beam is 8 keV. A synthesized strain field monotonically decreases from top-right to bottom-left; meanwhile, the composition *x* of the $Al_xGa_{1\text{-}x}N$ linearly increases from 0.13 to 0.18 along the X-direction. Three circular regions with almost uniform strain distribution are artificially inserted at positions (*X, Y*) = (9.5, 9.5), (15.5, 25.5), (23.5, 12.5) μm, with radii equal to 2 μm, 4 μm, and 6 μm, respectively, as shown in Fig. 2(a). For the ease of simulation, it is assumed that the XRD profile from positions with a large strain has a large FWHM and low peak height, as illustrated in Fig. 2(b) and (c). Representative simulated XRD profiles are listed in Fig. 2(e). Additional details about the synthesized dataset are provided in ***Supplementary Note 3***.

Briefly, the similarity score calculated from the *neighborhood-based similarity metric* measures the difference between two sets of neighbors N_i^S and N_i^D. The calculated distribution of similarity scores is listed in Fig. 2(e) and (f), corresponding to $amp = 0$ and $amp = 1.1$, respectively. The similarity score is 1 for all sampling points when $amp = 0$ (Fig. 2(e)) since there are no XRD features in *D* $(X,\ Y,\ amp{\cdot}nor\backslash\_XRD)$. On the other hand, Fig. 2(f) shows that the low-score area clearly corresponds to the boundary assumed in the synthetic dataset, while the score is insensitive to the continuous background strain gradient. The results in Fig. 2 demonstrate the effectiveness of the *neighborhood-based similarity metric*, particularly its ability to identify regions with structural continuity changes in the material, thereby providing valuable information for subsequent analysis.

3.2. Case study I: Bulk GaN crystal with multiple domains grown by FFC Na-flux method

We further analyze the practical nanoXRD results obtained from a cross-sectional GaN wafer grown by the Na-flux method with flux film coating technique (FFC), as an example of a crystal containing multiple domains. Details of the sample growth can be found in Refs.34, 37–39. Fig. 3 shows the cross-sectional multiphoton excitation photoluminescence microscopy (MPPL) image of the sample, where several clear boundaries of different growth sectors marked by b1 to b5 (dashed black lines) are formed. On this crystal, due to the varying growth facets and oxygen impurity incorporation that occurred during the growth process, multiple growth areas appear as different photoluminescence intensities and patterns; b1 locates within the *c*-plane growth sector, b2 and b4 are boundaries between $\{1\bar{1}01\}$ and $\{1\bar{1}02\}$ facets growth sectors, b3 is the boundary between c-plane and $\{1\bar{1}02\}$ facets growth sectors, b5 is the boundary between the left and right parts of $\{1\bar{1}01\}$ and $\{1\bar{1}02\}$ facets growth sectors.

The measurement area, marked by the red rectangle in Fig. 3, is 120 × 300 μm^2, with a sampling interval of 5 μm. The incident angle of the nanoXRD beam (*ω*) scans with a step of 0.002° within a range of 0.2°.

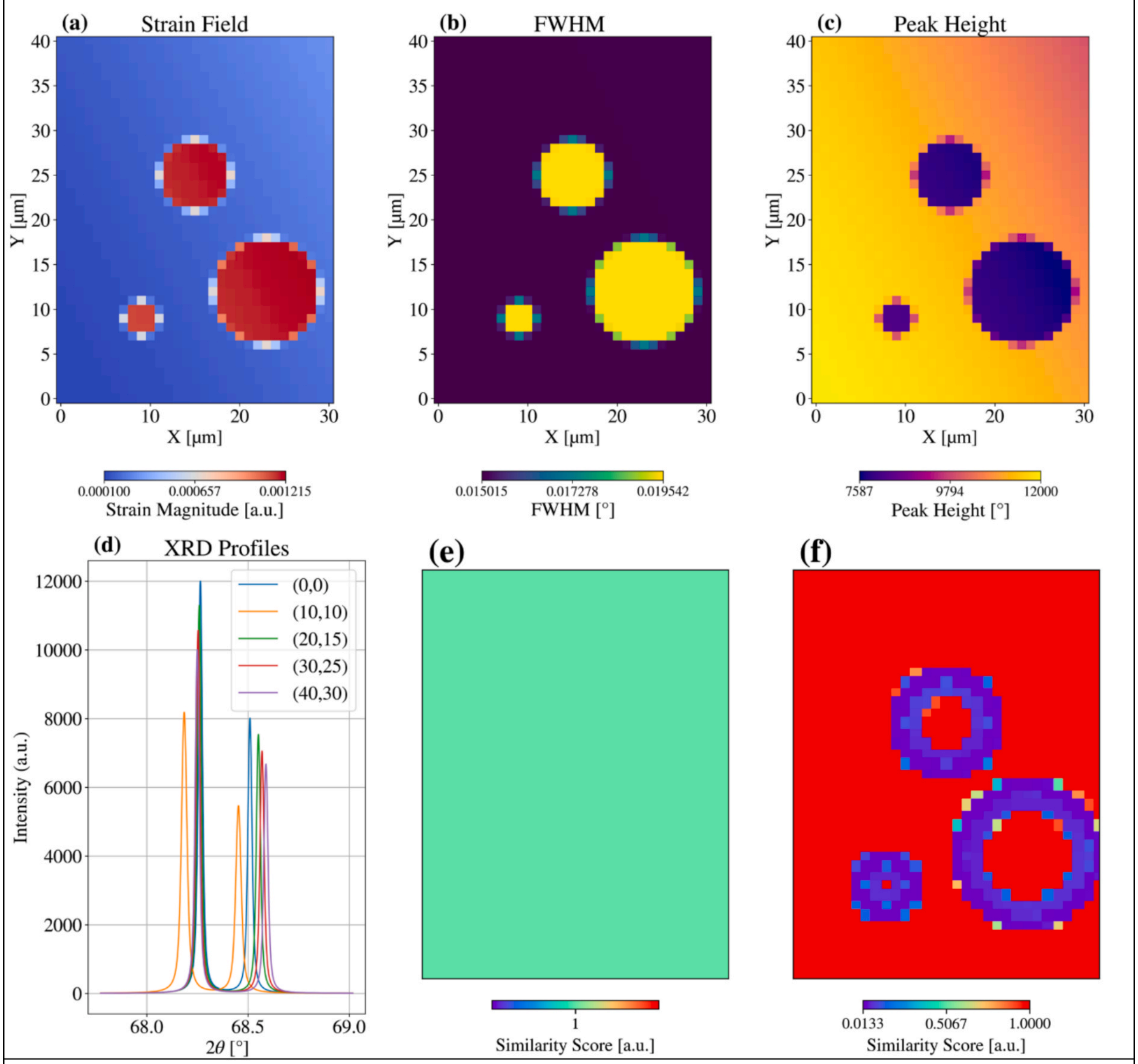


Fig. 2. Illustration of the synthesized dataset. The 2D distribution maps of (a) strain field, (b) FWHM, and (c) Peak heights. (d) Representative XRD spectrum of points located at positions of $(X, Y) = (0, 0), (10, 10), (20, 15), (30, 25)$, and $(40, 30)$ µm. The similarity score J distribution is mapped when (e) *amp* = 0, (f) *amp* = 1.1.

For each sampling point, symmetric $(2\bar{2}00)$ and asymmetric $(2\bar{2}02)$ planes of GaN were chosen as diffraction planes.

In analyzing the experimental nanoXRD data, the *neighborhood-based similarity metric* was first used to obtain the score distribution with *amp* = 13 for $2\bar{2}00$ reflections, *amp* = 15 for $2\bar{2}02$ reflections, as shown in Fig. 4(a) and (c). Based on the score distribution map, image segmentation was then performed to divide the original data into several regions, as shown in Fig. 4(b) and (d) using a watershed-based approach [40,41] implemented with the *OpenCV* library (OpenCV's marker-based watershed). Further computational details, including the stability of similarity score maps with respect to *amp* variation and the image segmentation procedure, are provided in ***Supplementary Notes 4*** and ***5***. In particular, ***Supplementary Note 4*** (**Figs. S5(h) and S6(h)**) demonstrates that both spatial and diffraction information in D are essential for quantifying the location of discontinuities, and that *amp* controls how strongly these discontinuities disrupt local continuity.

Boundaries b1 to b5 observed in MPPL are plotted as dashed black lines in Fig. 4. Besides, boundaries of segmented areas are labeled with −1, and the background area is labeled with 1. It is also worth noting that, in some instances of watershed calculation, labels associated with small regions are sometimes overwritten by the expansion of neighboring labels representing larger areas, resulting in the disappearance of labels from small regions in the final watershed result. In our case, this effect caused regions 11 in Figs. 4(b) and 8 in Fig. 4(d) to vanish.

Box plots of the similarity scores for each labeled area are shown in Fig. 5 to highlight areas with potentially significant structural variations. Based on the relationship between the similarity score and structural variations, greater structural variations in a region result in a lower score distribution, which appears as a lower value in the box plot. In the following section, regions 7, 12, and 10 in Fig. 4(b)

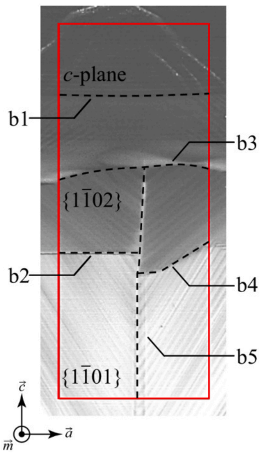


Fig. 3. Cross-sectional MPPL image of the GaN sample grown by the FFC Na-flux method.

(corresponding to regions 7, 12, and 11 in Fig. 4(d)) are selected for detailed discussion. For clarity, these regions are hereafter referred to as regions A, B, and C, respectively. The selection of regions A, B, and C is motivated by their distinct structural characteristics: regions A and B exhibit the lowest similarity score distributions (Fig. 5), whereas distributions and shapes of regions A and B show little connection to the known boundaries (b1–b5). Meanwhile, region C is directly related to b2, b4, and b5 but exhibits comparatively higher score distribution in Fig. 5.

Angular characteristics of diffraction profiles provide key information for clarifying the nature of structural deformations. For instance, an increased density of lattice defects broadens the diffraction peaks, resulting in a broader peak; impurity incorporation alters the local interplanar spacings, leading to detectable shifts in the 2θ peak positions. Furthermore, fluctuations in the morphology of lattice planes manifest as changes in the tilting angle ω. Peak positions in 1D XRD profiles are therefore essential for identifying structural similarities and differences among regions A, B, and C. However, it has been reported that in practical nanoXRD data analysis, XRD profiles are often asymmetric, making fitting with traditional Gaussian or Voigt functions difficult [33]. To eliminate uncertainties arising from unreliable fitting, we employed a fitting-free statistical approach, as described in more detail in ***Supplementary Note 6***. In short, the method characterizes the peak positions of each angular diffraction intensity profile by calculating the relative diffraction peak centroids with respect to the averaged profile, rather than relying on the fitted parameters from traditional Gaussian or Voigt functions.

Using the calculated relative diffraction peak positions, that is, centroids of ω and 2θ, we evaluated the structural similarities and differences among regions A, B, and C, as shown in Figs. 6 and 7. Measurement points are indicated by black and red dots, with the red dots located at the boundaries between each labeled area. It is also worth noting that $2\bar{2}02$ reflections contain contributions from both *m*-and *c*-planes, while $2\bar{2}00$ reflections reflect only the characteristics of *m*-planes Fig. 8.

Fig. 6**(a–1), (a–2), and** Fig. 7**(a–1), (a–2)** reveal that both regions A and B exhibit highly localized structural deformations associated with the $2\bar{2}00$ reflections, whose positions are marked by the white rectangle and arrows. Fig. 6 (ω centroid map) shows that extreme maxima and minima of tilt appear around these localized deformation points, while Fig. 7 (2θ centroid map) further demonstrates that corresponding extrema in interplanar spacing are also observed in the same regions.

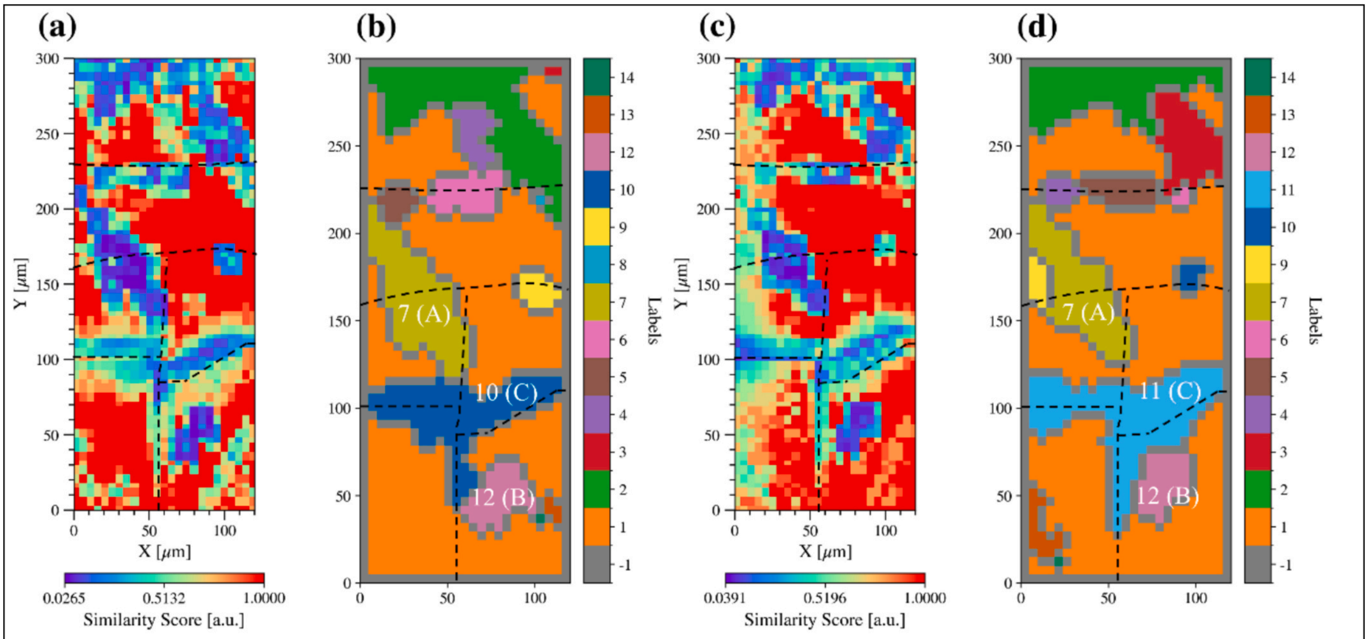


Fig. 4. Analysis results of practical nanoXRD measurement. Results from the $2\bar{2}00$ reflections are shown for (a) the similarity score distribution, and (b) the segmented image. Results from the $2\bar{2}02$ reflections are shown for (c) the similarity score distribution and (d) the segmented image.

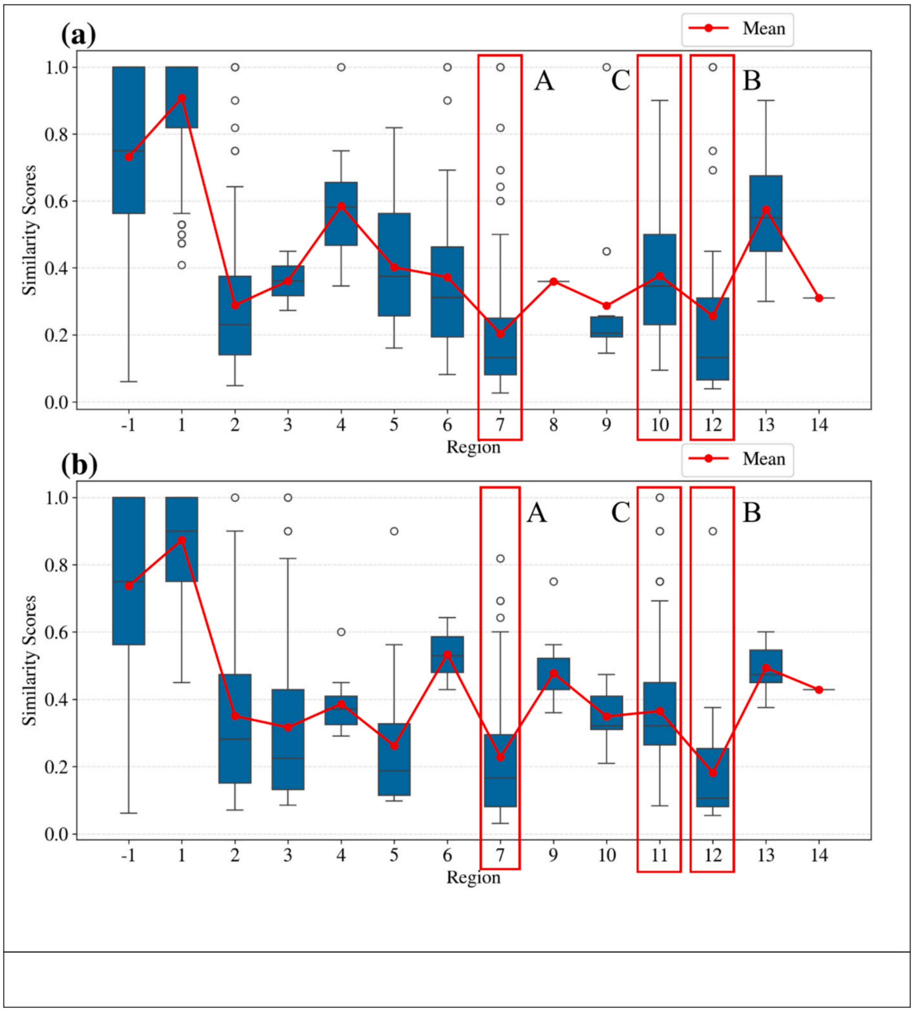


Fig. 5. Box plots of the region-dependent similarity score distribution. Results are shown for (a) $2\bar{2}00$ reflections and (b) $2\bar{2}02$ reflections.

Similar behavior is found for the $2\bar{2}02$ reflections.

Unlike regions A and B, on the other hand, region C, corresponding to b2, b4, and b5, exhibits localized but milder structural deformation extending over a broader area with a narrower variation range in the color-map scales of Fig. 6**(a–3), (b–3) and** Fig. 7**(a–3), (b–3)**. We find that b2 and b4, which lie at the boundaries between $\{1\bar{1}02\}$ and $\{1\bar{1}01\}$ growth sectors, tend to show maxima in ω and a sharp increase in 2θ centroids along the Y-direction. In addition, the structural variation associated with b5 (indicated by red arrows in Fig. 6**(a–3), (b–3) and** Fig. 7**(a–3), (b–3)**) appears slightly to the right of the b5 position observed in the MPPL image. This discrepancy likely arises because the b5 position observed in Fig. 3 is determined from MPPL measurements of the sample surface, whereas nanoXRD probes deeper into the crystal, leading to a shift away from b5.

In regions A and B, the highly localized structural deformations suggest that end-on dislocation cores are the primary defect sources [10]. To strengthen the physical interpretation of the low-similarity-score regions associated with dislocations, we introduced an independent, strain-based validation in ***Supplementary Note 7***. The qualitative agreement between the similarity metric and the theoretical strain-field model demonstrates that the proposed analysis framework not only enables data-driven detection of structural discontinuities but also provides physically meaningful insight into variations of the crystal structure. While higher spatial-resolution measurements would be required for a more definitive, defect-specific interpretation, **Ref.** [10] reports that dislocation-induced effects can be detected by nanoXRD at distances of 5 μm from the dislocation core.

In contrast, region C exhibits a different type of structural variation unrelated to dislocations. Oxygen is known to be the most common unintentionally incorporated impurity (UII) in FFC GaN, with its concentration varying across different growth sectors [37–39]. Previous studies have also shown that variations in UII oxygen concentration across growth sectors result in luminescence contrast in MPPL images [42,43]. Therefore, the sharp contrast across b2 and b4 in Fig. 3 is likely caused by such variations in UII oxygen concentration.

The resulting oxygen-induced lattice mismatch between the $\{1\bar{1}01\}$ and $\{1\bar{1}02\}$ growth sectors is reflected in an increase in the 2θ centroids across b2 and b4 (Fig. 7**(a–3) and (b–3)**). To accommodate this mismatch, a transition region is expected to form, as schematically illustrated in Fig. 8, and is characterized by morphological variations. This region exhibits an increase in the tilting angle ω for both *m*- and *c*-planes, consistent with the trends observed in Fig. 6**(a–3) and (b–3)**. Notably, Fig. 6**(a–3)** shows a larger tilt at b2, while Fig. 6**(b–3)** indicates a larger tilt at b4. Given that $2\bar{2}02$ reflections include contributions from *c*-planes in addition to *m*-planes, the difference between Fig. 6**(a–3) and (b–3)** suggests that *c*-planes have a greater tilting than *m*-planes at b4, whereas the opposite trend occurs at b2.

Except for b2 and b4, b5 corresponds to the sample's coalescence boundary (CB) [34,37–39]. The extremely low 2θ values observed in Fig. 7**(a–3) and (b–3)** are likely due to localized stress concentration at this boundary and to oxygen impurity segregation, both of which contribute to an enlarged lattice spacing. Notably, the extreme-value

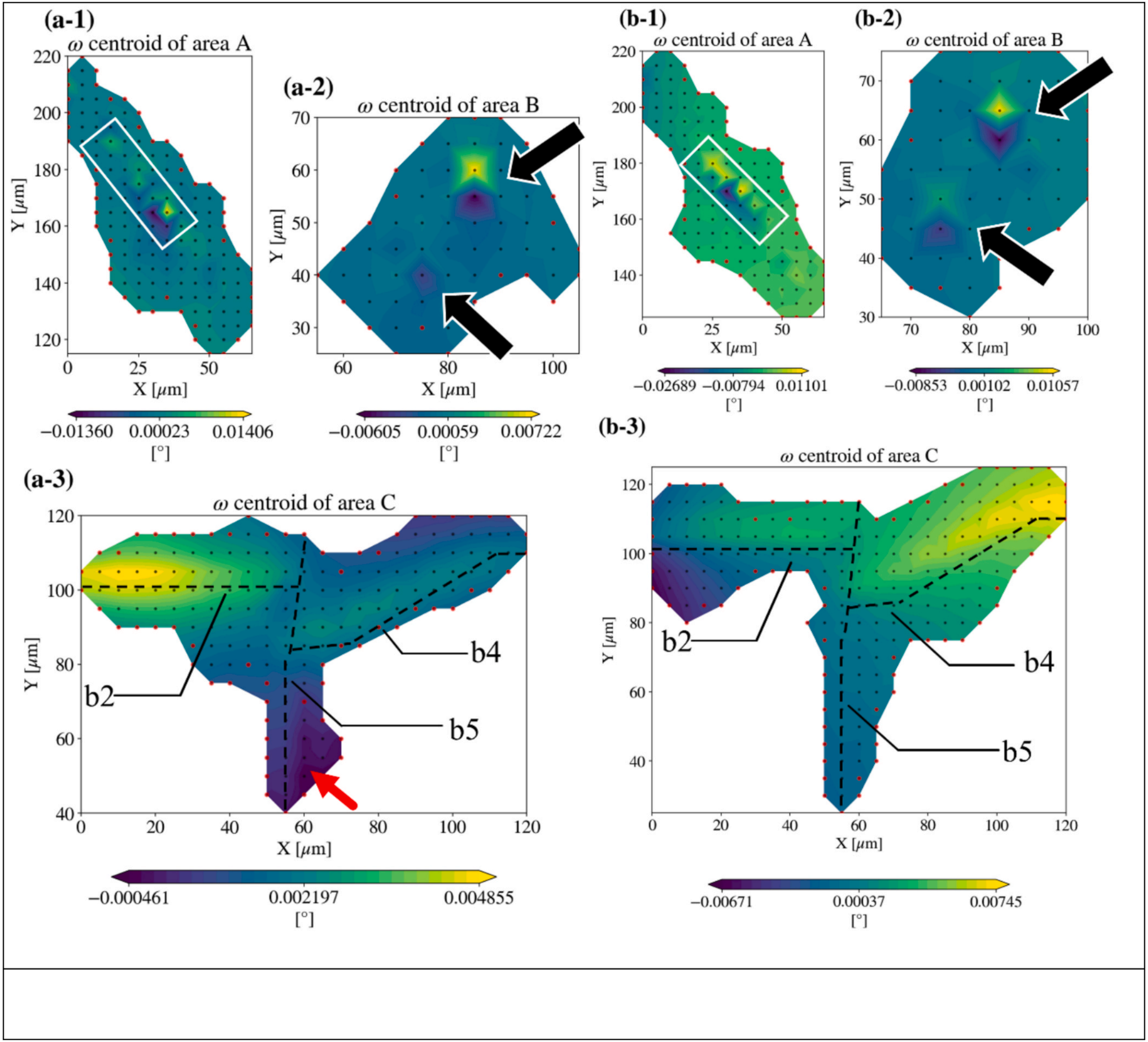


Fig. 6. Maps of ω centroids. Results for (a-1) region A, (a-2) region B, and (a-3) region C in $2\bar{2}00$ reflections and for (b-1) region A, (b-2) region B, and (b-3) region C in $2\bar{2}02$ reflections are shown.

distribution of 2θ is confined to the CB within the $\{1\bar{1}01\}$ growth sector, indicating that coalescence-induced stress is also restricted to this sector. In contrast, among the results in Fig. 6, only Fig. 6**(a–3)** shows a decrease in ω toward b5 (indicated by the red arrow), suggesting that the lattice tilt induced by coalescence is more pronounced along the *m*-plane than along the *c*-planes.

Collectively, these results demonstrate that nanoXRD effectively captures impurity-induced boundaries between the $\{1\bar{1}02\}$ and $\{1\bar{1}01\}$ growth sectors and the existence of localized defects, presumably attributed to dislocations. By resolving full features of planes within the diffraction volume, nanoXRD provides a comprehensive picture of how dislocations, impurity inhomogeneities, and crystal coalescence generate structural discontinuities, thereby extending the insights gained from MPPL imaging. Additional changes in diffraction peaks in regions A, B, and C are detailed in ***Supplementary Note 8***. We further verified similar behavior in additional regions of the FFC GaN sample, as presented in ***Supplementary Note 9***, supporting the conclusion that the observed variations in oxygen concentration are not attributable to data collection or processing artifacts.

3.3. *Case study II: Bulk GaN crystal with few boundaries grown by OVPE and HVPE*

To further validate the performance of the *neighborhood-based similarity metric* on samples with fewer domains, boundaries, or interfaces, nanoXRD experiments were performed on an OVPE GaN sample and an HVPE sample, as shown in Figs. 9 and 10.

As discussed in Refs.44–47, the growth of OVPE samples typically produces numerous inverted pyramidal pits, characteristic of a 3D growth mode. These pyramidal growth pits consist of $\{11\bar{2}2\}$ and $\{3\bar{3}04\}$ facets, which are highlighted in grey and red, respectively, as illustrated in Fig. 9(a). In the MPPL images from the *c*-plane and cross-sectional *m*-plane shown in Fig. 9(b) and 9(c), $\{11\bar{2}2\}$ and $\{3\bar{3}04\}$ facets appear as regions of dark and bright contrast, respectively,

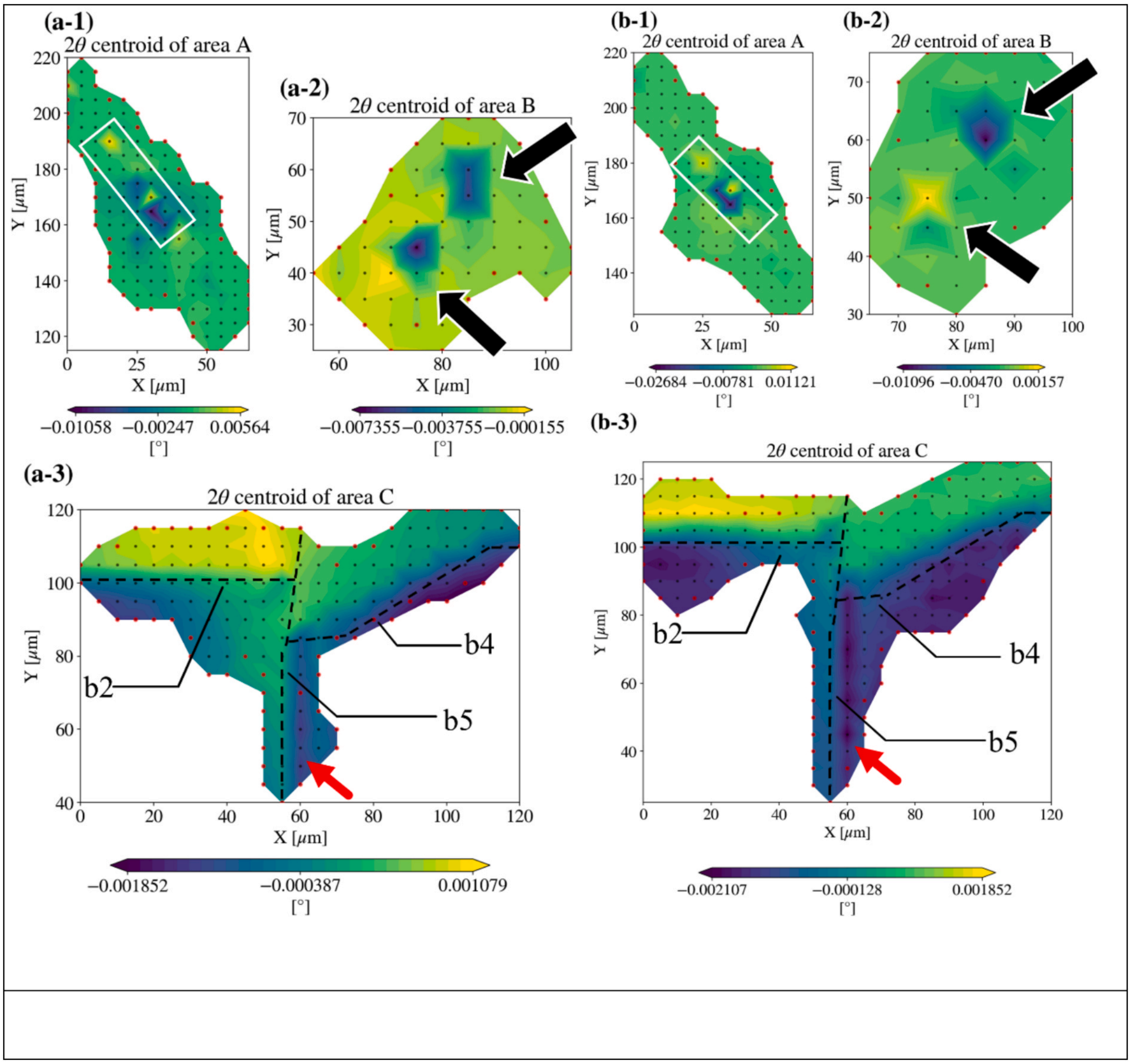


Fig. 7. Maps of 2θ centroids. Results for (a-1) region A, (a-2) region B, and (a-3) region C in $2\bar{2}00$ reflections and for (b-1) region A, (b-2) region B, and (b-3) region C in $2\bar{2}02$ reflections are shown.

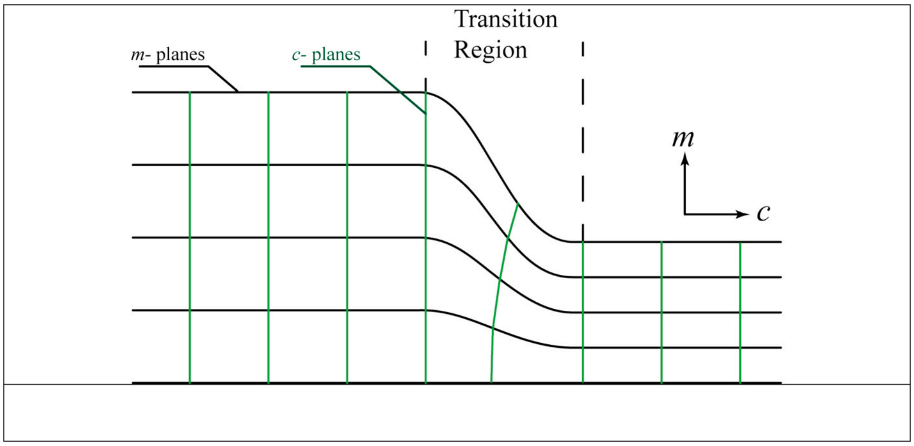


Fig. 8. Schematic illustration of the transition region. The figure is drawn in an exaggerated manner.

corresponding to different concentrations of incorporated oxygen impurity. Earlier studies in Refs.45–47 indicate that threading

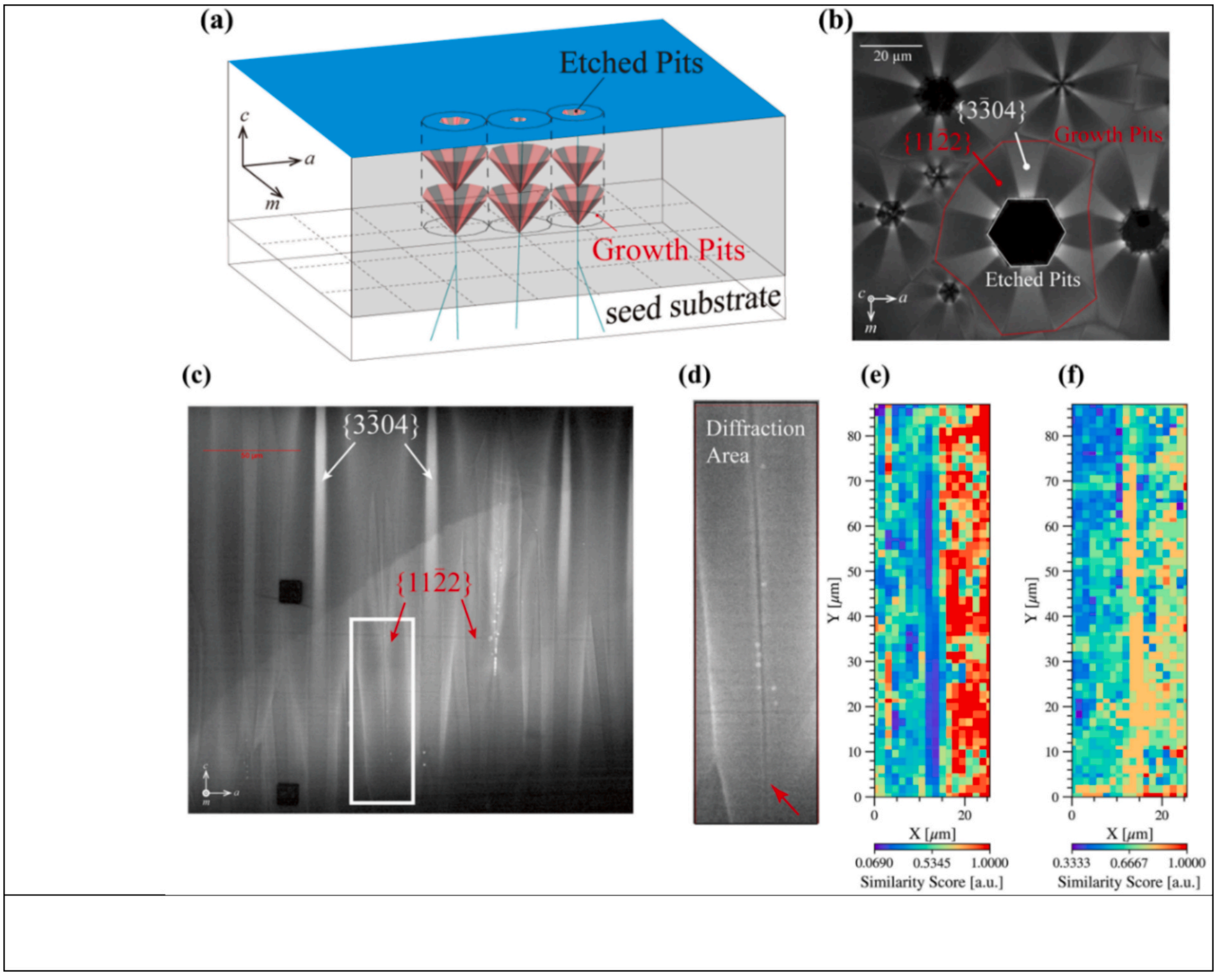


Fig. 9. Results of an OVPE GaN sample. (a) Schematic of the crystal structure in the OVPE sample. (b) The MPPL image of the *c*-plane of the sample. (c) Cross-sectional *m*-plane's MPPL image of the sample. The white rectangle marks the area analyzed by nanoXRD. (d) Magnified MPPL image of the analyzed area. The red arrow indicates the center of the growth pit. Maps of similarity score distribution obtained from the *neighborhood-based similarity metric* are shown for (e) $2\bar{2}00$ reflections and (f) $2\bar{2}02$ reflections

dislocations (TDs) tend to converge at centers of pyramidal pits, a feature confirmed by the positions of the TD-related etch pits formed using a NaOH-KOH solution, as shown in Fig. 9(b). A magnified cross-sectional image near the center of a growth pit reveals a stem-like contrast (red arrow in Fig. 9(d)). Considering that the dislocation core acts as a non-radiative recombination center [48], this stem-like contrast is attributed to the presence of TDs.

The position-dependent nanoXRD measurements using $2\bar{2}00$ and $2\bar{2}02$ reflections were conducted on the area shown in Fig. 9(d) which has a size of 25.5 × 87 μm^2, with an interval of 1.5 μm between sampling points. The incident angle of the nanoXRD beam scans with a step of 0.002° within a range of 0.2°. The score distribution maps of the *neighborhood-based similarity metric* are depicted in Fig. 9(e) and (f) for $2\bar{2}00$ and $2\bar{2}02$ reflections, respectively. In Fig. 9(e), low similarity scores are distributed along the Y (or *c*-axis) direction at the center of the image, whereas Fig. 9(f) shows high scores in the same region. Both patterns are attributed to the presence of a TD. The results from the $2\bar{2}00$ reflections (Fig. 9(e)) show low crystal structural continuity near the center, while the $2\bar{2}02$ reflections (Fig. 9(f)) exhibit relatively high continuity at the same location. The difference at the center of the measurement area can be explained by the contributions from the *c*-planes, which are included only in the $2\bar{2}02$ reflections. Comparing Fig. 9(e) and (f) suggests that th. This implies that the TD is likely of edge type, since the Burgers vector of an edge TD lies in the basal plane with minimal impact on the *c*-plane. Besides, both figures consistently show higher similarity scores on the right side of the TD compared with the left, indicating structural asymmetry within the pyramidal growth pit. These results demonstrate that the *neighborhood-based similarity metric*, combining nanoXRD measurements, effectively captures plane-dependent variations in crystal structure with TDs.

The second example involves a previously reported HVPE GaN sample, with detailed nanoXRD experimental procedures described in **Ref.** [33]. Although the MPPL image in Fig. 10(a) indicates a pronounced structural boundary near Y = 17 μm, earlier work showed that the $2\bar{2}00$ and $2\bar{2}02$ reflections exhibit distinct structural features across this boundary. In particular, these reflections together revealed the presence of two growth boundaries located at $Y_1 = 17$ μm and $Y_2 = 27$ μm [33]. The results in Fig. 10(b) and (c) confirm this previous analysis and, in addition, provide further insights into the associated structural transitions. Compared with the bottom part of the sample, the $2\bar{2}02$ reflection exhibits a decline in continuity around Y_1, reaching a minimum near Y_2. This distribution agrees with our earlier conclusion that misfit occurs near Y_1 and that stacking faults are generated around Y_2. Interestingly, although our earlier study reported severe broadening of nanoXRD diffraction profiles caused by stacking faults for Y > 27 μm, the results in Fig. 10(b) and (c) show that structural continuity is well preserved on the *c*-plane (Fig. 10(c)) but strongly disrupted on the *m*-plane (Fig. 10(b)) for the region at Y > 27 μm, X > 20 μm. This difference between *c*-plane and *m*-plane again highlights the anisotropy of

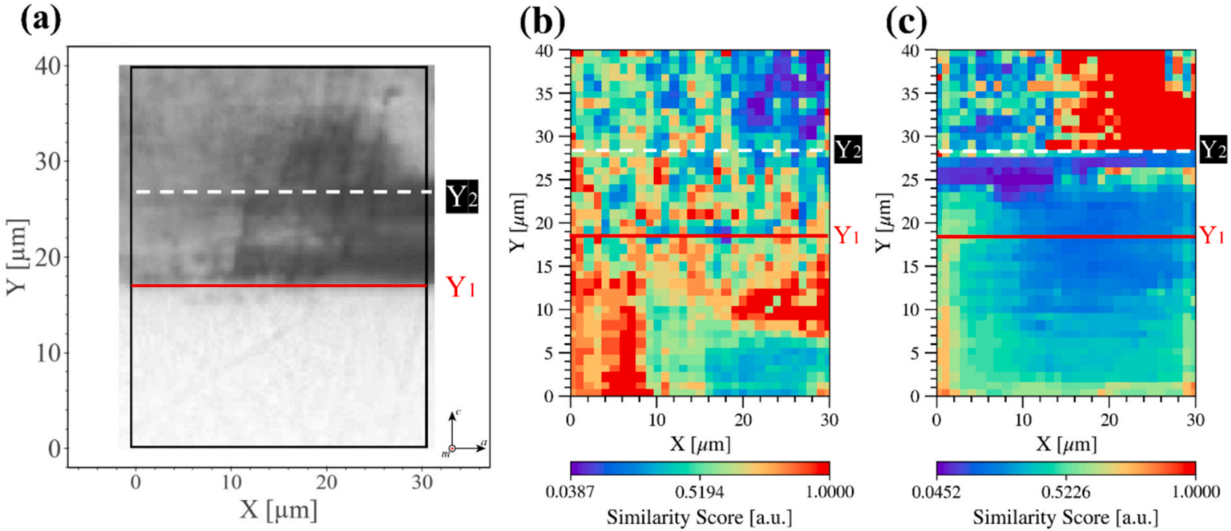


Fig. 10. Results for an HVPE GaN sample. (a) Cross-sectional MPPL image of the sample. Maps of similarity score distribution from the *neighborhood-based similarity metric* are shown for (b) $2\bar{2}00$ reflections and (c) $2\bar{2}02$ reflections.

structural deformation in regions affected by stacking fault evolution.

3.4. Physical interpretability, constraints, and insights of the metric

We also note that interpreting the outputs of exploratory, unsupervised analysis inevitably requires additional physical insight, particularly when identifying interfaces or defects such as dislocations. One route toward a more definitive interpretation is to increase the spatial resolution of nanoXRD measurements, guided by regions of interest highlighted by the similarity score maps. Another important route is validation through independent, physics-based modeling or simulations. In this work, we introduced a strain-field-based simulation of a screw dislocation in ***Supplementary Note 7***. The qualitative agreement between the similarity metric and the strain-field model provides independent validation that the proposed framework not only enables data-driven discovery of structural discontinuities, but also supports physically meaningful interpretation of the crystal structure.

We further note that the locality enforced by the proposed similarity metric is physically consistent with post-growth nanoXRD analysis of cross-sectional samples. While similar diffraction patterns may arise at spatially non-neighboring locations on topmost crystal surfaces due to quasi-two-dimensional growth, such behavior was not observed in the cross-sectional datasets analyzed here. Along the growth direction, structural variations continuously evolve in the absence of abrupt changes in growth conditions, making the occurrence of identical diffraction patterns at distant locations unlikely. By explicitly incorporating spatial coordinates into the data space, the proposed framework enforces a physically motivated locality constraint that suppresses non-local similarities, ensuring that reduced similarity scores reflect local disruptions in structural continuity rather than global repetition. Consistent with this interpretation, additional analysis in ***Supplementary Note 4*** shows that the parameter *amp* effectively regularizes the data structure by controlling the strength of this locality constraint.

Beyond structural segmentation, the proposed framework enables new insights into the relationship between crystal processing and microstructural defect formation. Because the same similarity metric can be applied consistently across different samples, reflections, and growth conditions, it provides a generalized basis for comparing how specific fabrication processes manifest as localized structural discontinuities. For example, in Fig. 10(c), the strongly disrupted continuity observed around region Y_2 suggests a localized structural response that is likely associated with changes in the HVPE growth process. Although the present work does not attempt to reconstruct the full growth history, such observations demonstrate how the method can systematically highlight regions where processing conditions may have induced distinct defect structures or strain states. In this sense, the approach serves as an exploratory but physically grounded tool for connecting fabrication conditions with resulting microstructural features.

4. Conclusion

In conclusion, the *neighborhood-based similarity metric* provides a general and interpretable framework for nanoXRD-based crystal structure analysis by integrating diffraction information with spatial coordinates. This approach enables the quantitative detection of structural transitions across growth-sector interfaces and domain boundaries that conventional clustering or peak-fitting methods may overlook. In the analyzed GaN crystals, distinct growth sectors exhibit well-separated regions of low similarity, facilitating the identification of interfacial boundaries and local strain variations **probably ignored in other observation methods**. Even in samples exhibiting fewer domains or more continuous structures, where segmentation becomes challenging, the spatial distribution of similarity scores still captures subtle structural gradients and transitions linked to different reflection planes.

Furthermore, the method supports a forward-looking workflow in which the analysis of nanoXRD datasets using this similarity metric guides the construction of defect-aware, process-linked reference datasets. By consistently detecting and localizing different types of structural discontinuities across samples grown under varied conditions, the framework provides a foundation for data-driven labeling and categorization of defects. Such datasets could subsequently enable more targeted supervised learning approaches, facilitating the detection of even subtle post-growth structural variations associated with specific growth conditions.

Finally, because the continuity evaluation relies on local neighborhood relationships rather than on data type or dimensionality, this framework is broadly applicable to other spatially or temporally resolved techniques, such as Raman spectroscopy or electron diffraction, especially in cases requiring the detection of weak or gradual structural transitions. Although the present study does not include *in situ* on-the-fly experiments, the same analysis concept can, in principle, be extended to such measurements. For example, replacing spatial coordinates in the joint data space with time- or temperature-dependent variables would enable the neighborhood-consistency framework to detect abrupt structural changes, phase transitions, or non-continuous structural responses as a function of time or temperature.

CRediT authorship contribution statement

Zhendong Wu: Writing – original draft, Visualization, Validation, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Tetsuya Tohei:** Investigation, Writing – review & editing, Supervision, Funding acquisition. **Yusuke Hayashi:** Investigation, Funding acquisition. **Shigeyoshi Usami:** Resources. **Masayuki Imanishi:** Resources. **Yusuke Mori:** Resources. **Junichi Takino:** Resources. **Kazushi Sumitani:** Resources. **Yasuhiko Imai:** Resources. **Shigeru Kimura:** Resources. **Akira Sakai:** Writing – review & editing, Supervision, Project administration, Funding acquisition, Conceptualization.

Funding

This work was partially supported by JSPS KAKENHI (JP16H06423 (S.A.), JP20H00352 (Y.M.), JP22KK0055 (Y.H.), JP23H01447 (Y.H.)) and Murata Science and Education Foundation (T.T.,Y.H.).

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

We thank Dr. B.Y. Zhang for the helpful discussion on the machine learning algorithm.

The synchrotron radiation experiments were performed at the BL13XU of SPring-8 with the approval of the Japan Synchrotron Radiation Research Institute (JASRI) (Proposal Nos. 2019B1797, 2020A1331, 2021A1584, 2021B1345, 2021B1650, 2022B1567, 2024A1637, 2024A1821, 2024A1929, 2024B1651, and 2025A1547).

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.matdes.2026.115669.

Data availability

I have shared the link to the code and data in the manuscript.

The codes and processed datasets used in this study are publicly available at https://github.com/aldnoahone/metrics_similarity.git. The raw nanoXRD diffraction data, exceeding 1 TB in total volume, are not included in this repository but can be obtained from the corresponding author upon reasonable request or through access to BL13XU at SPring-8.

References

[1] C.E. Murray, A. Ying, S.M. Polvino, I.C. Noyan, M. Holt, J. Maser, Nanoscale silicon-on-insulator deformation induced by stressed liner structures, J. Appl. Phys. 109 (2011) 083543.

[2] T. Thorbeck, N.M. Zimmerman, Formation of strain-induced quantum dots in gated semiconductor nanostructures, AIP Adv. 5 (2015) 087107.

[3] K. Shida, S. Takeuchi, Y. Imai, S. Kimura, A. Schulze, M. Caymax, A. Sakai, Tomographic mapping analysis in the depth direction of high-Ge-content SiGe layers with compositionally graded buffers using nanobeam X-ray diffraction, ACS Appl. Mater. Interfaces 9 (2017) 13726–13732.

[4] K. Shida, N. Yamamoto, T. Tohei, M. Imanishi, Y. Mori, K. Sumitani, Y. Imai, S. Kimura, A. Sakai, Quantitative analysis of lattice plane microstructure in the growth direction of a modified Na-flux GaN crystal using nanobeam X-ray diffraction, Jpn. J. Appl. Phys. 58 (2019) SCCB16.

[5] K. Shida, S. Takeuchi, T. Tohei, H. Miyake, K. Hiramatsu, K. Sumitani, Y. Imai, S. Kimura, A. Sakai, Microstructural analysis in the depth direction of a heteroepitaxial AlN thick film grown on a trench-patterned template by nanobeam X-ray diffraction, J. Appl. Phys. 123 (2018) 161563.

[6] K. Shida, S. Takeuchi, T. Tohei, Y. Imai, S. Kimura, A. Schulze, M. Caymax, A. Sakai, Depth-resolved analysis of lattice distortions in high-Ge-content SiGe/compositionally graded SiGe films using nanobeam x-ray diffraction, Semicond. Sci. Technol. 33 (2018) 124005.

[7] T. Uchiyama, S. Takeuchi, S. Kamada, T. Arauchi, Y. Hashimoto, K. Yamane, N. Okada, Y. Imai, S. Kimura, K. Tadatomo, A. Sakai, Positional dependence of defect distribution in semipolar hydride vapor phase epitaxy-GaN films grown on patterned sapphire substrates, Jpn. J. Appl. Phys. 55 (2016) 05FA07.

[8] S. Kamada, S. Takeuchi, D.T. Khan, H. Miyake, K. Hiramatsu, Y. Imai, S. Kimura, A. Sakai, Microstructural analysis of an epitaxial AlN thick film/trench-patterned template by three-dimensional reciprocal lattice space mapping technique, Appl. Phys Express 9 (2016) 111001.

[9] Y. Imai, K. Sumitani, S. Kimura, Current status of nanobeam x-ray diffraction station at SPring-8, AIP Conference Proceedings 2054 (2019) 050004; https://doi.org/10.1063/1.5084622.

[10] T. Hamachi, T. Tohei, Y. Hayashi, S. Usami, M. Imanishi, Y. Mori, K. Sumitani, Y. Imai, S. Kimura, A. Sakai, Analysis of local strain fields around individual threading dislocations in GaN substrates by nanobeam x-ray diffraction, J. Appl. Phys. 135 (2024) 225702.

[11] S. Aoyagi, A. Aoyagi, H. Takeda, H. Osawa, K. Sumitani, Y. Imai, S. Kimura, Time-resolved nanobeam X-ray diffraction of a relaxor ferroelectric single crystal under an alternating electric field, Crystals 11 (2021) 1419.

[12] S. Aoyagi, A. Aoyagi, H. Takeda, H. Osawa, K. Sumitani, Y. Imai, S. Kimura, Position and electric field dependent local lattice strain detected by nanobeam x-ray diffraction on a relaxor ferroelectric single crystal, Phys. Rev. B./phys. Rev. B 105 (2022) 024101.

[13] H. Shiomi, A. Ueda, T. Tohei, Y. Imai, T. Hamachi, K. Sumitani, S. Kimura, Y. Ando, T. Hashizume, A. Sakai, Analysis of inverse-piezoelectric-effect-induced lattice deformation in AlGaN/GaN high-electron-mobility transistors by time-resolved synchrotron radiation nanobeam X-ray diffraction, Appl. Phys Express 14 (2021) 095502.

[14] A. Shimada, H. Shiomi, T. Tohei, Y. Hayashi, M. Yamaguchi, J. Yamamoto, T. Hamachi, Y. Imai, K. Sumitani, S. Kimura, S. Kaneki, T. Hashizume, A. Sakai, In situ nanobeam x-ray diffraction of local strain in AlGaN/GaN high-electron-mobility transistors under operating condition, J. Appl. Phys. 138 (2025) 075701.

[15] N. Hrauda, J. Zhang, E. Wintersberger, T. Etzelstorfer, B. Mandl, J. Stangl, D. Carbone, V. Holý, V. Jovanović, C. Biasotto, L.K. Nanver, J. Moers, D. Grützmacher, G. Bauer, X-ray nanodiffraction on a single SiGe quantum dot inside a functioning field-effect transistor, Nano Lett. 11 (2011) 2875–2880.

[16] M.C. Marshall, D.F. Phillips, M.J. Turner, M.J.H. Ku, T. Zhou, N. Delegan, F. J. Heremans, M.V. Holt, R.L. Walsworth, Scanning X-Ray diffraction microscopy for diamond quantum sensing, Phys. Rev. Appl 16 (2021) 054032.

[17] T.U. Schülli, S.J. Leake, X-ray nanobeam diffraction imaging of materials, Curr. Opin. Solid State Mater. Sci. 22 (2018) 188–201.

[18] C. Corley-Wiciak, M.H. Zoellner, I. Zaitsev, K. Anand, E. Zatterin, Y. Yamamoto, A. A. Corley-Wiciak, F. Reichmann, W. Langheinrich, L.R. Schreiber, C.L. Manganelli, M. Virgilio, C. Richter, G. Capellini, Lattice deformation at submicron scale: X-Ray nanobeam measurements of elastic strain in electron shuttling devices, Phys. Rev. Appl 20 (2023) 024056.

[19] C. Corley-Wiciak, M.H. Zoellner, A.A. Corley-Wiciak, F. Rovaris, E. Zatterin, I. Zaitsev, G. Sfuncia, G. Nicotra, D. Spirito, N. Von Den Driesch, C.L. Manganelli, A. Marzegalli, T.U. Schulli, D. Buca, F. Montalenti, G. Capellini, C. Richter, Full picture of lattice deformation in a $Ge_{1-x}Sn_x$ Micro-Disk by 5D X-ray diffraction microscopy, Small Methods 8 (2024) 2400598.

[20] A. Luo, T. Zhou, M.V. Holt, A. Singer, M.J. Cherukara, Deep learning of structural morphology imaged by scanning X-ray diffraction microscopy, Sci. Rep. 15 (2025) 21736.

[21] V.-A. Surdu, R. Győrgy, X-ray diffraction data analysis by machine learning methods—a review, Appl. Sci. 13 (2023) 9992.

[22] Z. Chen, N. Andrejevic, N.C. Drucker, T. Nguyen, R.P. Xian, T. Smidt, Y. Wang, R. Ernstorfer, D.A. Tennant, M. Chan, M. Li, Machine learning on neutron and x-ray scattering and spectroscopies, Chem. Phys. Rev. 2 (2021) 031301.

[23] G.-X. Zhang, Y. Song, W. Zhao, H. An, J. Wang, Machine learning-facilitated multiscale imaging for energy materials, Cell Rep. Phys. Sci. 3 (2022) 101008.

[24] K. Choudhary, B. DeCost, C. Chen, A. Jain, F. Tavazza, R. Cohn, C.W. Park, A. Choudhary, A. Agrawal, S.J.L. Billinge, E. Holm, S.P. Ong, C. Wolverton, Recent advances and applications of deep learning methods in materials science, npj Comput. Mater. 8 (2022) 59.

[25] C. Wang, U. Steiner, A. Sepe, Synchrotron big data science, Small 14 (2018) 1802291.

[26] Y. Sun, S. Brockhauser, P. Hegedűs, C. Plückthun, L. Gelisio, D.E.F. De Lima, Application of self-supervised approaches to the classification of X-ray diffraction spectra during phase transitions, Sci. Rep. 13 (2023) 9370.

[27] P.M. Vecsei, K. Choo, J. Chang, T. Neupert, Neural network based classification of crystal symmetries from x-ray diffraction patterns, Phys. Rev. B 99 (2019) 245120.

[28] M. Kodama, A. Ohashi, H. Adachi, T. Miyuki, A. Takeuchi, M. Yasutake, K. Uesugi, T. Kaburagi, S. Hirai, Three-dimensional structural measurement and material identification of an all-solid-state lithium-ion battery by X-Ray nanotomography and deep learning, J. Power Sources Adv. 8 (2021) 100048.

[29] H. Chan, Y.S.G. Nashed, S. Kandel, S.O. Hruszkewycz, S.K.R.S. Sankaranarayanan, R.J. Harder, M.J. Cherukara, Rapid 3D nanoscale coherent imaging via physics-aware deep learning, Appl. Phys. Rev. 8 (2021) 021407.

[30] K.G. Yager, P.W. Majewski, M.M. Noack, M. Fukuto, Autonomous x-ray scattering, Nanotechnology 34 (2023) 322001.

[31] A.E. Siemenn, E. Aissi, F. Sheng, A. Tiihonen, H. Kavak, B. Das, T. Buonassisi, Using scalable computer vision to automate high-throughput semiconductor characterization, Nat. Commun. 15 (2024) 4654.

[32] P.M. Maffettone, A.C. Daly, D. Olds, Constrained non-negative matrix factorization enabling real-time insights of in situ and high-throughput experiments, Appl. Phys. Rev. 8 (2021) 041410.

[33] Z. Wu, Y. Hayashi, T. Tohei, K. Sumitani, Y. Imai, S. Kimura, A. Sakai, Machine learning assisted nanobeam X-ray diffraction based analysis on hydride vapor-phase epitaxy GaN, J. Appl. Cryst. 58 (2025) 1205–1219.

[34] Z. Wu, T. Tohei, M. Imanishi, Y. Mori, K. Sumitani, Y. Imai, S. Kimura, A. Sakai, Revealing crystallographically distinct growth sectors in Na-flux GaN via machine learning-assisted nanobeam x-ray diffraction analysis, APL Mach. Learn. 3 (2025) 036114.

[35] L. McInnes, J. Healy, N. Saul, L. Großberger, UMAP: uniform manifold approximation and projection, J. Open Source Softw. 3 (2018) 861.

[36] L. Van Der Maaten, G. Hinton, Visualizing Data using t-SNE, Journal of Machine Learning Research 9 (2008) 2579–2605. Website: https://www.jmlr.org/papers/v9/vandermaaten08a.html.

[37] M. Imanishi, K. Murakami, T. Yamada, K. Kakinouchi, K. Nakamura, T. Kitamura, K. Okumura, M. Yoshimura, Y. Mori, Promotion of lateral growth of GaN crystals on point seeds by extraction of substrates from melt in the Na-flux method, Appl. Phys Express 12 (2019) 045508.

[38] Y. Mori, M. Imanishi, K. Murakami, M. Yoshimura, Recent progress of Na-flux method for GaN crystal growth, Jpn. J. Appl. Phys. 58 (2019) SC0803.

[39] M. Imanishi, K. Okumura, K. Nakamura, T. Kitamura, K. Kakinouchi, K. Murakami, M. Yoshimura, Y. Fujita, Y. Tsusaka, J. Matsui, Y. Mori, Anomalous dislocation annihilation behavior observed in a GaN crystal grown on point seeds by the Na-flux method, Appl. Phys Express 13 (2020) 085510.

[40] A.S. Kornilov, I.V. Safonov, An overview of watershed algorithm implementations in open source libraries, J. Imaging 4 (2018) 123.

[41] A. Kornilov, I. Safonov, I. Yakimchuk, A review of watershed implementations for segmentation of volumetric images, J. Imaging 8 (2022) 127.

[42] T. Hamachi, T. Tohei, M. Imanishi, Y. Mori, A. Sakai, Correlation between current leakage and structural properties of threading dislocations in GaN bulk single crystals grown using a Na-flux method, Jpn. J. Appl. Phys. 58 (2019) SCCB23.

[43] T. Hamachi, T. Tohei, M. Imanishi, Y. Mori, A. Sakai, Local current leakage at threading dislocations in GaN bulk single crystals grown by a modified Na-flux method, Jpn. J. Appl. Phys. 58 (2019) 050918.

[44] T. Ueda, GaN power devices: current status and future challenges, Jpn. J. Appl. Phys. 58 (2019) SC0804.

[45] J. Takino, T. Sumi, Y. Okayama, M. Nobuoka, A. Kitamoto, M. Imanishi, M. Yoshimura, Y. Mori, Development of a 2-inch GaN wafer by using the oxide vapor phase epitaxy method, Jpn. J. Appl. Phys. 58 (2019) SC1043.

[46] J. Takino, T. Sumi, Y. Okayama, A. Kitamoto, S. Usami, M. Imanishi, M. Yoshimura, Y. Mori, Floral design GaN crystals: low-resistive and low-dislocation-density growth by oxide vapor phase epitaxy, Jpn. J. Appl. Phys. 60 (2021) 095501.

[47] J. Takino, T. Sumi, Y. Okayama, A. Kitamoto, M. Imanishi, M. Yoshimura, N. Asai, H. Ohta, T. Mishima, Y. Mori, Extreme reduction of on-resistance in vertical GaN p–n diodes by low dislocation density and high carrier concentration GaN wafers fabricated using oxide vapor phase epitaxy method, Appl. Phys Express 13 (2020) 071010.

[48] T. Sugahara, H. Sato, M. Hao, Y. Naoi, S. Kurai, S. Tottori, K. Yamashita, K. Nishino, L.T. Romano, S. Sakai, Direct evidence that dislocations are non-radiative recombination centers in GaN, Jpn. J. Appl. Phys. 37 (1998) L398.