

Article

Application of Machine Learning Methods to Neutron Transmission Spectroscopic Imaging for Solid–Liquid Phase Fraction Analysis

Takashi Kamiyama ^{1,*}, Kazuma Hirano ¹, Hiroataka Sato ¹, Kanta Ono ², Yuta Suzuki ², Daisuke Ito ³
and Yasushi Saito ³

- ¹ Graduate School of Engineering, Hokkaido University, Kita-13 Nishi-8, Kita-ku, Sapporo 060-8628, Japan; main.horse.hirano@gmail.com (K.H.); h.sato@eng.hokudai.ac.jp (H.S.)
- ² Institute of Materials Structure Science, High Energy Accelerator Research Organization (KEK), Tsukuba 305-0801, Japan; kanta.ono@kek.jp (K.O.); ysuz@post.kek.jp (Y.S.)
- ³ Institute for Integrated Radiation and Nuclear Science, Kyoto University, Kumatori-cho, Osaka 590-0494, Japan; ito.daisuke.5a@kyoto-u.ac.jp (D.I.); ysaito@rri.kyoto-u.ac.jp (Y.S.)
- * Correspondence: takashik@eng.hokudai.ac.jp; Tel.: +81-11-706-6651

Abstract: In neutron transmission spectroscopic imaging, the transmission spectrum of each pixel on a two-dimensional detector is analyzed and the real-space distribution of microscopic information in an object is visualized with a wide field of view by mapping the obtained parameters. In the analysis of the transmission spectrum, since the spectrum can be classified with certain characteristics, it is possible for machine learning methods to be applied. In this study, we selected the subject of solid–liquid phase fraction imaging as the simplest application of the machine learning method. Firstly, liquid and solid transmission spectra have characteristic shapes, so spectrum classification according to their fraction can be carried out. Unsupervised and supervised machine learning analysis methods were tested and evaluated with simulated datasets of solid–liquid spectrum combinations. Then, the established methods were used to perform an analysis with actual measured spectrum datasets. As a result, the solid–liquid interface zone was specified from the solid–liquid phase fraction imaging using machine learning analysis.

Keywords: neutron transmission spectroscopic imaging; machine learning analysis; solid–liquid phase fraction; solidification process; lead–bismuth eutectic alloy



Citation: Kamiyama, T.; Hirano, K.; Sato, H.; Ono, K.; Suzuki, Y.; Ito, D.; Saito, Y. Application of Machine Learning Methods to Neutron Transmission Spectroscopic Imaging for Solid–Liquid Phase Fraction Analysis. *Appl. Sci.* **2021**, *11*, 5988. <https://doi.org/10.3390/app11135988>

Academic Editor: Wolfgang Treimer

Received: 28 May 2021
Accepted: 23 June 2021
Published: 27 June 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

In neutron transmission spectroscopic imaging (NTSI), microscopic information is mapped on an object image with a wide field of view by analyzing the transmission spectrum of each point in two-dimensional imaging data [1–5]. Since the shape of the transmission spectrum is determined by the neutron total cross-section, which reflects the microstructure of the object material, it is possible to visualize the spatial distributions of analyzed microstructural parameters [6]. Since the distribution of statistically averaged microstructure information deep inside the sample can be visualized with a wide field of view, its application has been expanded beyond material analysis to include analysis of the deformation behavior of components [7] and evaluation of the homogeneity of actual products [8].

In recent years, machine learning (ML) has been used for various spectral analyses, including neutron science [9–16]. In neutron transmission spectrum analysis, which is the basis of NTSI method, it is considered that ML can be applied if the obtained spectrum can be categorized. That is, spectra with various material parameters are learned by ML in advance and the material parameters of the experimental spectra are determined using the ML model. When ML is used, it is not necessary to perform a detailed time-consuming

analysis for each pixel, so it can be expected that the efficiency will be greatly improved, especially in the case of spectroscopic/energy-dependent imaging with a significant number of pixels.

However, the issue in this case is the collection of a large number of spectra with a wide variety of microscopic structures needed to build the ML model. In other words, it is necessary to collect a number of Bragg-edge spectrum patterns of the solid material with various different microstructures, which is the main subject in NTSI analysis. Bragg-edge patterns reflecting various material microstructures can be calculated, for example, using RITS (Rietveld Imaging of Transmission Spectra) code [6]. However, it is difficult to calculate all types of microstructures as training spectra; moreover, there are patterns that cannot be calculated with analysis codes such as RITS. Therefore, we considered building an ML model for experimental analysis using the experimentally measured Bragg-edge spectra as training data for the ML.

Our choice as the first trial for such ML analysis was the imaging of solid–liquid interfaces in motion. When the solid–liquid interface in the solidification process is measured by neutron imaging, due to the relationship between the measurement time and the solidification velocity, a zone appears where the solid and liquid phases coexist between the 100% solid-phase and 100% liquid-phase regions. In this study, we applied ML to the solid–liquid phase fraction analysis of this solid–liquid coexisting zone. In this case, the analysis was not used for the microstructure of the solid phase itself directly, so it was easy to apply ML. In addition, various Bragg-edge data can be collected using the solid-phase spectrum of the entire solidified sample. From the above, it is considered that the analysis of the solid–liquid phase fraction is optimal for evaluating the applicability of the ML analysis method to NTSI as a first step.

In this study, lead-bismuth eutectic alloy (LBE), which our group has experience in analyzing [17], was used as a sample for imaging the solidification process. LBE is expected to be used as a coolant and spallation target for next-generation nuclear reactors (accelerator-driven subcritical reactors), and it is required for evaluating the heat transfer coefficient by determining the moving velocity of the solid–liquid interface [18–20]. As it is difficult to analyze the solid–liquid phase fraction of neutron transmission spectra with high accuracy for LBE due to its complicated crystal phase [21], this is considered an appropriate application of ML analysis.

Therefore, the aim of this study was to investigate the possibility of applying ML analysis to neutron transmission spectroscopy through LBE solid–liquid phase fraction imaging as an example. For the ML analysis, the solid–liquid phase fraction of each pixel during solidification was evaluated by the ML model built from the training dataset, created from 100% liquid-phase and 100% solid-phase transmission measurement data. We first evaluated the analytical performance of the ML using the neutron transmission spectral dataset generated by the neutron total cross-section simulation. Then, ML was applied to the experimental transmission imaging data for the solidification process of LBE, which solidification process was performed to simulate the coolant solidification event, and then, solid–liquid phase fraction imaging was performed.

2. Overview and Outline of Machine Learning Analysis

2.1. Application of Machine Learning Methods to Solid–Liquid Phase Fraction Analysis

In this study, to learn the spectra for LBE solids with various crystalline textures produced by solidification, we assembled training data from the solidified LBE data itself at the time of the experiment. In addition, the neutron transmission spectrum handled here was composed of a large number of data points by the neutron time-of-flight (TOF) method, and the computational cost was too large to be applied to ML in its original form. Therefore, we first reduced the dimensionality of the spectral data through unsupervised machine learning and then used the reduced dimensionality data to discriminate the spectral shape by supervised machine learning in the second stage. The procedure for the analysis was as follows:

- (1) Measure the solidification process of liquid LBE with NTSI.
- (2a) Obtain the solid-phase spectrum and liquid-phase spectrum of LBE from the imaging data. For the solid-phase spectrum, after the whole sample has solidified, divide the solid zone into several parts and obtain multiple spectra with various crystalline textures from each part. For the liquid-phase spectrum, the neutron data for the entire melted liquid zone are integrated to obtain a liquid-state spectrum.
- (2b) Create a training dataset with different crystalline textures and solid–liquid fractions by adding the liquid-phase spectrum to each solid-phase spectrum acquired in step (2a) in fractions from 0 to 100%.
- (2c) Reduce the dimensionality of the training data in step (2b) using unsupervised machine learning, then build an ML model using the training dataset through supervised machine learning.
- (3) Reduce the dimensionality of the experimental spectra of step (1) using unsupervised ML.
- (4) Apply the supervised machine learning model built in step (2c) to the dimensionality-reduced experimental spectra in step (3) and obtain the solid–liquid phase fraction. Visualize the obtained phase fraction at each pixel.

In this study, we built ML programs with python and scikit-learn libraries, which are commonly used.

2.2. Unsupervised Machine Learning

The purpose of the dimensionality reduction of high-dimensional data such as neutron transmission spectra can be classified into two main categories. The first of these is data compression. This can reduce computational costs when using supervised ML, where datasets tend to be large. The other purpose is data visualization. This can help to convert high-dimensional data whose characteristics are difficult to recognize into a two- or three-dimensional diagram. For these purposes, unsupervised ML can be used as a method to perform the dimensionality reduction of high-dimensional data.

There are various methods for the dimensionality reduction of data using unsupervised ML, including principal component analysis (PCA), latent semantic analysis (LSA), nonnegative matrix factorization (NMF), latent Dirichlet allocation (LDA), etc. Among them, PCA is one of the most typical dimensionality reduction methods; it has a long history and is widely used in various situations [22,23]. There are two main methods used for the dimensionality reduction: selecting only important variables and not using the remaining variables, and building new variables from the variables of the original data. PCA applies to the latter. That is, original data can be explained by lower-dimensional variables converting high-dimensional data by PCA. In this study, we adopted PCA as a dimensionality reduction method for neutron transmission spectra.

In PCA, the following procedure is used to obtain the principal components.

- (1) Calculate the variance-covariance matrix.
- (2) Solve the eigenvalue problem for the variance-covariance matrix to find the eigenvectors and eigenvalues.
- (3) Represent data in the direction of each principal component.

Consider a variance-covariance matrix A for the given data:

$$Ax = \lambda x \quad (1)$$

where A is an n th-order square matrix and x is an n -dimensional vector. Finding x and the eigenvalue λ that satisfy (1) is mathematically equivalent to the problem of finding orthogonal axes that maximize the variance. In PCA, when the eigenvalue problem is solved for the variance-covariance matrix A , multiple combinations of eigenvalues and eigenvectors are calculated. The eigenvector with the largest eigenvalue is called the first principal component, while the eigenvector with the second largest eigenvalue is called the second principal component. After the third principal component, the eigenvectors

with the largest eigenvalues are expressed in order. Further, by dividing the eigenvalues calculated for each principal component by the sum of the eigenvalues, the importance of the principal components can be expressed as a ratio. This is called the contribution ratio, and it indicates how much each principal component expresses the data. The sum of the contribution rates in order from the first principal component is called the cumulative contribution rate, and this can be used to measure the amount of information lost through dimensionality reduction. The cumulative contribution rate is discussed in Section 3.2.

2.3. Supervised Machine Learning

There are several supervised ML methods [24] used to carry out regression analysis on dimensionally reduced data to perform spectral analysis, such as the K -nearest neighbor (KNN) algorithm, extra tree regression (ETR), and support vector regression (SVR). In this study, we adopted the KNN algorithm, which requires a lower cost of computation. KNN is a supervised ML method that can be used for both classification and regression problems. Supervised ML basically consists of two parts: the training part, which is used to calculate the optimal parameters from the training data, and the prediction part, which makes predictions using the calculated training parameters. However, KNN does not have the part to calculate the optimum parameters from the training data. Since the training data are learned and predictions are made directly, the computational load is small.

KNN is a simple algorithm that learns the training data they are when training. The procedure used for solving a regression problem for input data is as follows.

- (1) Calculate the distance between the input data and the training data.
- (2) Chose K training data from their data points closest to the input data.
- (3) Obtain the average value of the indications of the training data and use it as the result of the regression problem.

In KNN, when predicting a regression problem, the distance between the input data and the training data is generally calculated using Euclidean distance. When the two points P and Q —defined by variables p_i and q_i , respectively—for which we want to find the distance are represented by $P(p_1, p_2, \dots, p_n)$ and $Q(q_1, q_2, \dots, q_n)$, the multidimensional Euclidean distance $d(P, Q)$ is given by the following equation:

$$d(P, Q) = \sqrt{(p_1 - q_1)^2 + (p_2 - q_2)^2 + \dots + (p_n - q_n)^2}. \quad (2)$$

When KNN handles a large amount of data, the prediction of unknown data takes a long time because KNN performs a neighborhood search on a large amount of training data. In addition, since $d(P, Q)$ is used to measure the distance between data, there appears to be only a small distance difference between data when the number of data dimensions is large, and it is often impossible to train with a high accuracy. Therefore, KNN is effective when the amount of data (spectra) is small or when the number of dimensions (channel number of time-of-flight spectrum) is small. The number of neighborhood training data points, K , to be acquired when making predictions with KNN is not determined within the inference or prediction, and K is called hyperparameter. A smaller K leads to a more complicated decision boundary, therefore overfitting is more likely to occur. Therefore, to improve the generalization performance, it is necessary to adjust an appropriate K using training data.

3. Examination of Machine Learning Analysis Using Simulation Spectra

3.1. Preparation of Simulation Spectra for Evaluation of Machine Learning Analysis Methods

Prior to the solid–liquid phase fraction analysis of the measured spectra, the validity of the aforementioned analysis procedure was confirmed by simulated LBE Bragg-edge transmission spectra prepared using computer simulation shown below. The simulated total cross-section spectrum of the LBE solid phase was calculated using the RITS code [6] with the systematic variation of the texture information. The parameters considered here were the direction of the crystal orientation vector (preferred orientation $\langle hkl \rangle$) and the

degree of development of the orientation texture (March–Dollase coefficient, MD). However, since the ratio of the β -phase is higher than that of the γ -phase in LBE [20], only the texture of the β -phase was considered here. Sixteen types of texture combinations were created between four types of preferred orientation vectors ($\beta<100>$, $\beta<002>$, $\beta<101>$, $\beta<102>$) parallel to the neutron transmission direction and four types of MD coefficients (MD = 0.6, 0.7, 0.8, 0.9). A completely isotropic solid-phase texture (MD = 1.0) was added and we prepared 17 types of LBE solid-phase spectra in total. The total cross-section spectrum of the LBE liquid phase was prepared from the neutron transmission spectroscopic imaging experiment (see Section 4.2). The simulated transmission spectra of various phase fractions of the LBE mixed solid–liquid phase were calculated from these cross-sections according to the next equation:

$$Tr(\lambda) = \exp \left[- \left\{ \Sigma_{\text{solid}}(\lambda)at + \Sigma_{\text{liquid}}(\lambda)(1 - a)t \right\} \right]. \quad (3)$$

Here, $Tr(\lambda)$ is the neutron transmission spectrum; $\Sigma_{\text{solid}}(\lambda)$ and $\Sigma_{\text{liquid}}(\lambda)$ are the macroscopic total cross-section of the LBE solid and liquid phases, respectively; and a is the fraction of the solid phase ($0 \leq a \leq 1$). t is the thickness of the entire sample, and here it was set to 10 mm according to the experiment. The solid-phase fraction a was changed from 0 to 100% in 1% steps, and one type of 100% liquid phase, which is common to the LBE system, was added to make a total of 1701 types of spectral datasets for evaluation. Figure 1 shows an example of the neutron transmission spectrum variation in an LBE solid–liquid mixture when the texture of the solid phase is isotropic. It can be seen that the Bragg-edges of the solid phase appear to be relatively clear against the 100% of the liquid phase. Thus, we prepared neutron transmission spectra of the LBE for various solid–liquid phase fractions (0 to 100%) and various textures (17 types as mentioned above) of the solid phase.

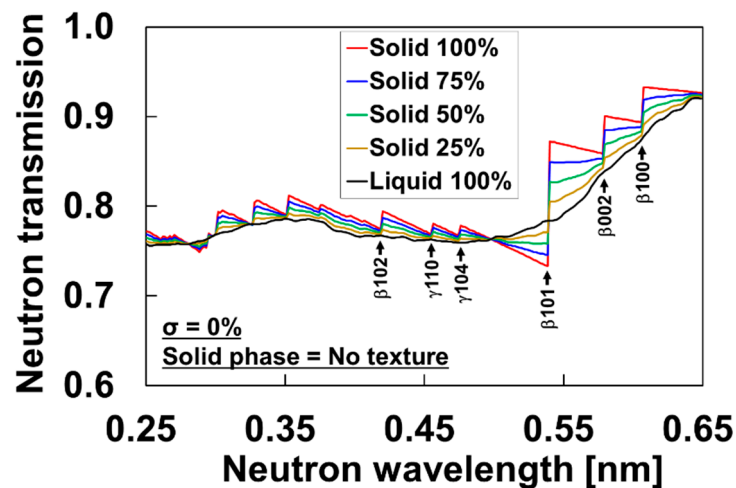


Figure 1. Neutron transmission spectra (Gaussian error $\sigma = 0\%$) of lead-bismuth eutectic alloy (LBE) simulating the solid–liquid mixed state when the solid phase had an isotropic texture (March–Dollase coefficient (MD) = 1.0).

Furthermore, in this study, to use the experimentally obtained transmission spectrum as training data, we also examined the case where the neutron transmission spectrum included statistical errors. To the simulated spectra created above, errors following a Gaussian distribution with a standard deviation σ of 1, 2, ..., 9, and 10% were randomly added to the simulated spectra. As an example, Figure 2 shows the neutron transmission spectra with $\sigma = 5\%$ error in the case of the isotropic solid-phase texture. From this, when the statistical error included in the neutron transmission spectrum becomes large, it becomes more difficult for human eyes to distinguish between the solid and liquid phases.

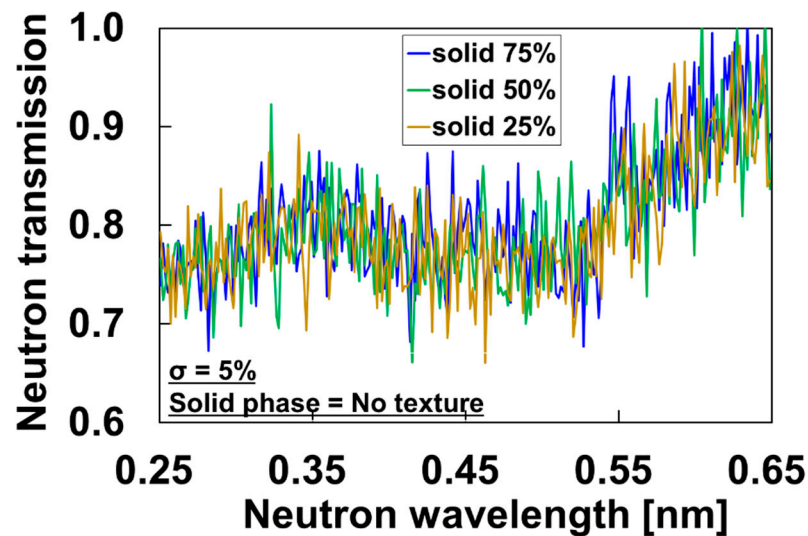


Figure 2. Neutron transmission spectra with random errors ($\sigma = 5\%$) simulating the solid-liquid mixed state when the solid phase had an isotropic texture ($MD = 1.0$).

3.2. Dimensionality Reduction Using Unsupervised Machine Learning Methods

Figure 3 shows the result of the cumulative contribution evaluation of the transmission spectrum dataset with $\sigma = 0\%$ created in Section 3.1 to evaluate the amount of information lost due to the dimensionality reduction of PCA. The horizontal axis of the figure is the number of principal components after dimensionality reduction by PCA, while the vertical axis is the cumulative contribution rate explained in Section 2.2. From the cumulative contribution rate, it is evident that more than 99% of the original data can be expressed with the first three principal components of the PCA; namely, this dataset can be expected to be sufficiently analyzed by reducing the number of dimensions to three.

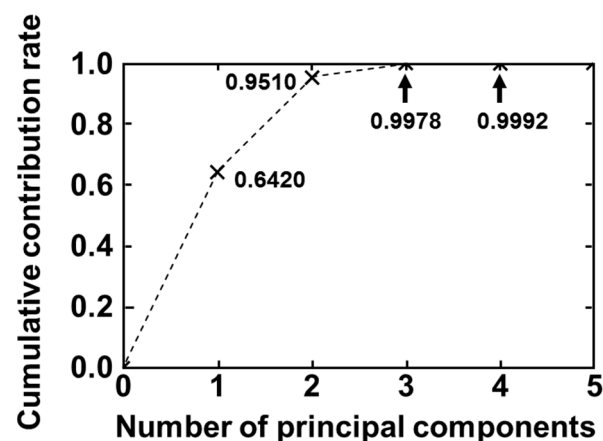


Figure 3. Cumulative rate of contribution to the number of dimension reductions (Gaussian error $\sigma = 0\%$).

Figure 4 shows the diagram plotting the distribution of the first three principal components of PCA expressing the simulated dataset of $\sigma = 0\%$. Note that each principal component is computationally found and the unit cannot be defined. In the figure, the point where each plot line converges is the point of a 100% liquid phase; from there, the point of a 100% solid phase appears at the opposite end of the line according to each texture state. Then, the data for each phase fraction are plotted on a straight line connecting both points. This straight line is drawn depending on the solid-liquid phase fraction. The data groups of different preferred orientation vectors are represented by colors, and the

black line in the center of the distribution corresponds to the solid–liquid phase fraction depending on the line of the isotropic texture. These groups are distributed to form fan-shaped planes in different preferred orientations from the black line. That is, as the texture develops (from MD = 1.0 to MD = 0.6), the data are plotted farther from the isotropic texture case (black line). In this way, it was found that by reducing the dimensions with PCA, it is possible to clearly identify the LBE transmission spectra with different solid–liquid phase fractions, preferred orientation vectors, and degrees of development of the texture. With this detailed analysis, it may be possible to obtain not only the solid–liquid phase fraction but also the texture information for the solid phase depending on the position of the transmission spectrum in this diagram.

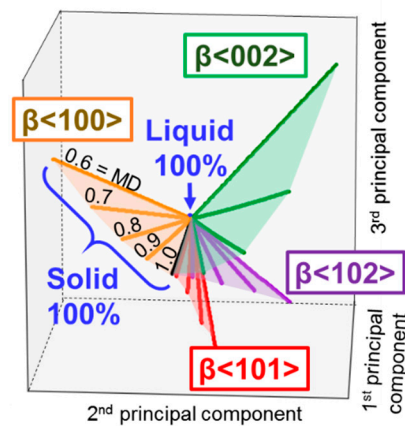


Figure 4. Representation of the simulated spectral dataset ($\sigma = 0\%$) dimensionally reduced to the first three principal components of principal component analysis (PCA).

3.3. Solid–Liquid Phase Fraction Analysis by Supervised Machine Learning Method

To evaluate the performance of the supervised machine learning method, the simulated dataset was randomly divided into training and testing data groups after dimensionality reduction by PCA. Of the original dataset including 1701 spectra, 1500 were for the training data group and 201 were for the test data group. The test data group was analyzed by the ML model built using the training data group, and the difference between the analyzed solid fraction and the original solid fraction was evaluated by the root mean square error (RMSE) shown by the equation:

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (f_i - y_i)^2}. \quad (4)$$

Here, n is the number of spectra analyzed, f_i is the solid-phase fraction estimated by the ML analysis, and y_i is the original solid-phase fraction set when the dataset was created. For the ML analysis, we tested SVR and ETR in addition to KNN. KNN and ETR showed higher performance than SVR when the training data contained statistical errors, and KNN had the lowest computational cost, so we adopted KNN in this study. The computational cost is also important in analyzing more than 10,000 spectra of all pixels of a two-dimensional detector in NTSI. In KNN, it is necessary to optimize the hyperparameter K of the training data, and in this study it was determined by 10-fold cross validation (CV).

First, the result of the solid-phase fraction analysis by KNN is shown in Figure 5. In this case, the Gaussian distribution error was not added to the simulated spectrum dataset ($\sigma = 0\%$). Here, $K = 2$ was used due to optimization. The horizontal axis of the diagram is the original solid-phase fraction, and the vertical axis is the solid-phase fraction estimated by KNN. The analyzed results of the test data shown by the red circle are in good agreement with the ideal values shown by the black line. In the figure, we also show the analysis results of the training data (blue triangles) in order to see how accurate the

analysis can be when analyzing the data used for training. Although this depends on the type of ML algorithm used, if there is a difference between the training and test data distributions there is a possibility that overlearning is occurring. In this case, there appears to be no problem.

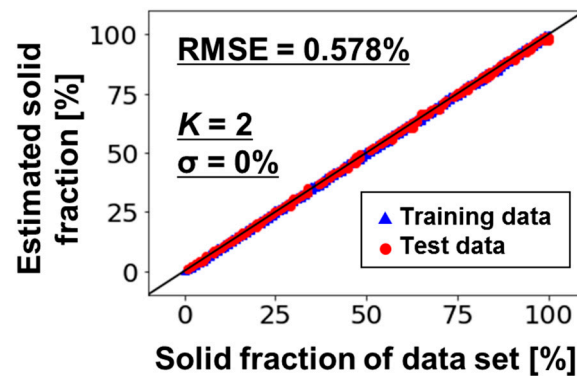


Figure 5. Results of solid–liquid phase fraction analysis by K -nearest neighbor algorithm (KNN) for the simulated spectral dataset without Gaussian errors ($\sigma = 0\%$) for both the training and testing data.

Next, we analyzed simulated datasets with Gaussian distribution errors ($\sigma = 1, 2, 3, \dots, 9, 10\%$) using the ML models built by the training spectra, including the errors, which is the subject of this research. For example, the ML analysis results for the dataset that included an error of $\sigma = 5\%$ were divided and analyzed as mentioned above and shown in Figure 6. Here, the hyperparameter $K = 10$ was the optimal value. As can be seen from the figure, the variance of both the training and test data groups increases around the black line, which is the set value of the solid-phase fraction. However, the estimated values follow the black line as a whole and the solid-phase fraction is considered to give results with a certain accuracy. Since the variances of the training and test data groups are about the same, note that our ML model did not fail with regard to overlearning.

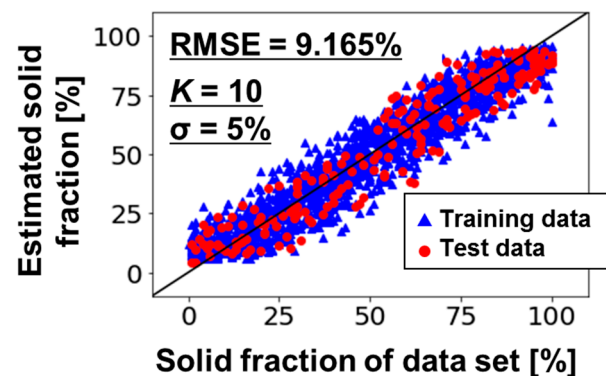


Figure 6. Results of solid–liquid phase fraction analysis by KNN for the simulated spectral dataset with random errors ($\sigma = 5\%$) for both the training and testing data.

Figure 7 shows the difference in RMSE when the ML models were built from each dataset created by changing the simulated statistical error, σ . The horizontal axis of the diagram is the magnitude of the statistical error σ in the test data and the vertical axis is the RMSE. This figure shows the difference in RMSE depending on the statistical error ($<5\%$) of the training data. If the statistical error contained in the test data is small, it is more accurate to perform the analysis using an ML model built using the training data of small σ . However, looking at more than 5% of the horizontal axis shown in the diagram (statistical error of the test (analyzed) data), we found that the accuracy of the analysis was almost the same even when the statistical error of the test data was larger than the error in

the training data. In other words, it is more profitable to obtain the training data with the same or better statistical accuracy than the analyzing data.

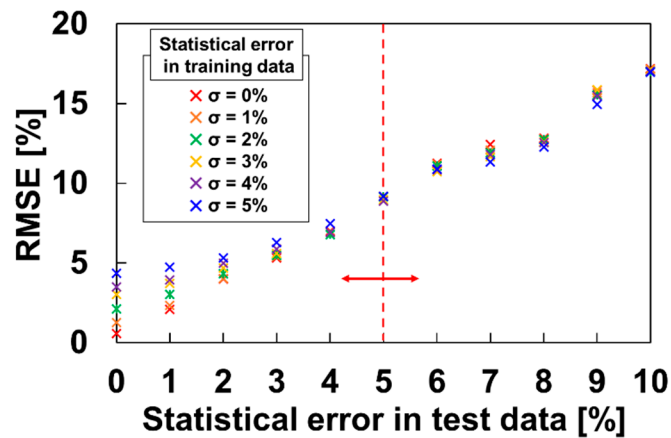


Figure 7. Difference in analytical performance (root mean square error (RMSE)) for the test data with random errors ($\sigma = 0–10\%$). The training data with random errors ($\sigma = 0–5\%$) were used to build the machine learning (ML) model using KNN.

4. Machine Learning Analysis Based on Actual Measurement Data

4.1. Overview of Neutron Transmission Spectroscopic Imaging Experiments of LBE Solidification

The NTSI experiments during the solidification process of LBE were performed using the energy-resolved neutron imaging system RADEN in the Materials and Life Science Experimental Facility (MLF) of the Japan Proton Accelerator Research Complex (J-PARC) [25] with the neutron wavelength resolution of $\approx 0.25\%$. The details of these experiments are already reported [26,27]. The LBE sample had a solidified volume 80 mm in vertical length, 80 mm in width, and 10 mm in thickness, and it was contained in a 304 stainless steel container with 1-mm-thick walls. The entire container was covered with glass fiber for heat insulation. A heater block at the top of the container and a heat sink at the bottom were installed. The sample temperature was checked using thermocouples inserted in seven locations around the container. A gas electron multiplier (GEM)-type detector (nGEM, Bee Beans Technologies [28]) was used as the neutron time-of-flight (TOF) imaging detector, covering a $100\text{ mm} \times 100\text{ mm}$ detection size. The pixel resolution was 0.8 mm in length and there were 128×128 pixels in total. In the experiment, the sample was heated for 60 min at first until about half of the top of the LBE sample reached a temperature higher than $124.7\text{ }^\circ\text{C}$, which is the melting point of LBE. Then, while holding the initial state, the first neutron transmission TOF imaging measurement was performed for 60 min. After that, the heating was stopped and the gradual solidification process from the downward direction of the sample was measured over 270 min as time-accumulated measurements every 30 min.

4.2. Creation of Training Data and Building Machine Learning Model

Figure 8 shows one of the neutron transmission images during the LBE solidification process as an example of the data measured. The figure shows a neutron wavelength-resolved radiogram with a neutron wavelength of $0.530–0.537\text{ nm}$ near the $\beta(101)$ Bragg-edge of LBE. The relatively uniform zone at the bottom of the figure is considered the zone where the sample did not melt during the first 60 min of heating. Upward from there, the zone where black and white vertical shades can be seen is considered the zone that was solidified through the solidification process. The slightly darker gray zone in the upper part (surrounded by a red frame) is estimated to be the zone where the solid–liquid interface moved during the 30 min integration measurement, and the uppermost uniform zone is the LBE melt. Although the approximate location of the solidification interface can be estimated

from the obtained images, it is necessary to evaluate the solid–liquid phase fraction to investigate its location in detail. The target of this study was to determine the rate at which this red zone moves from the bottom to the top during the solidification experiment.

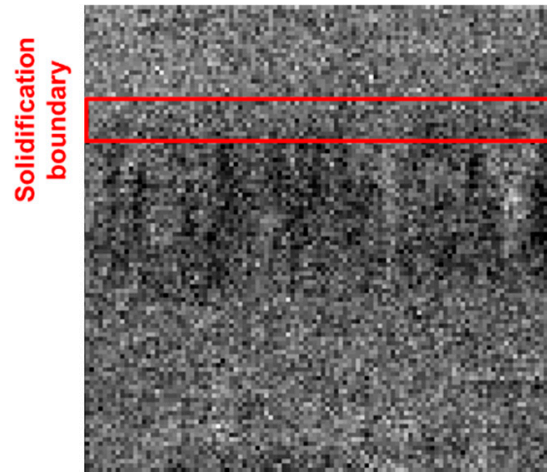


Figure 8. Neutron transmission imaging ($80\text{ mm} \times 80\text{ mm}$) of the LBE sample during solidification (neutron wavelength: $0.530\text{--}0.537\text{ nm}$; measurement time: 30 min). The red frame in the figure was considered the solid–liquid interface zone.

The training spectral dataset used for the ML analysis was created from neutron transmission spectra from TOF imaging experiments. For obtaining the liquid-phase spectrum shown in Figure 9, we averaged the entire zone where the liquid-phase was judged to be 100% in the first 60-min measurement, identified by the temperature information with the thermocouples inserted in the sample. For the spectrum of the solid phase, the data after the entire sample was solidified were picked out in a $45\text{ horizontal pixel} \times 2\text{ vertical pixel}$ section, avoiding the thermocouple insertion area, and these pixels were integrated and averaged together. Then, eventually, 50 types of solid-phase spectra were sampled for every two vertical pixels. Figure 10 shows some of the spectra obtained. The solid-phase spectra obtained from the different points show differences in texture. The effect of the sample container was removed as a background during these procedures.

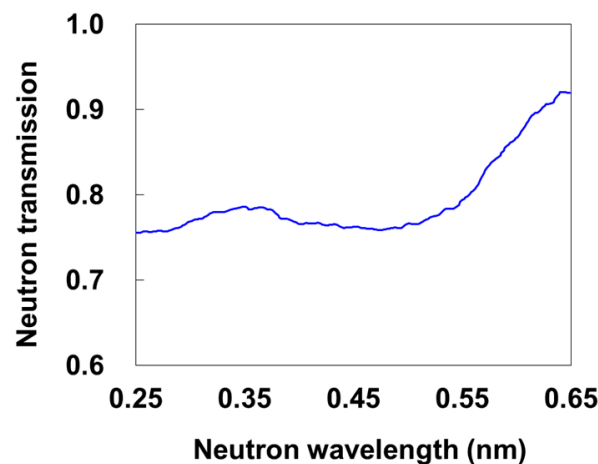


Figure 9. Neutron transmission spectrum of LBE melt obtained by integrating the 100% liquid-phase zone from the TOF imaging experiment.

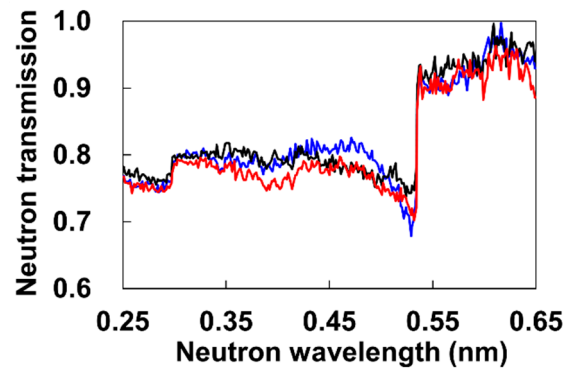


Figure 10. Example of neutron transmission spectra of 100% solid-phase LBE obtained from different points. The colors of the spectra are just for the visibility.

Using the spectra for the 100% liquid phase (1 kind) and 100% solid phase (50 kinds), prepared as described above, we created a training dataset of LBE neutron transmission spectra for solid-phase fractions from 0 to 100% in 1% increments following the procedure described in Section 3.1. Since we prepared 50 types of solid-phase spectra, the total number of neutron transmission spectra in the created training dataset was 5001. These spectra were then dimensionally reduced to 3 by PCA, and 5001 of these spectra were trained to build an ML model for the KNN.

4.3. Principal Component Imaging by PCA of Solidified Sample

As an example of the PCA result of the LBE spectra in the solidified sample obtained in the experiment, the first through fifth principal components are shown with the neutron transmission image in Figure 11. Different images were obtained for each principal component. The first principal component closely resembles the neutron transmission image, while the images of the second through fourth principal components are images that emphasize the solidification texture. It can also be seen that the higher the principal component is, the more information is lost in the image. The image shows an almost uniform noise distribution in the fifth principal component. Although these images do not directly correspond to physical quantities, PCA can visualize points of interest in the shape of the neutron transmission spectrum. Note that the sample container and thermocouples are clearly visible in the lower-order principal components.

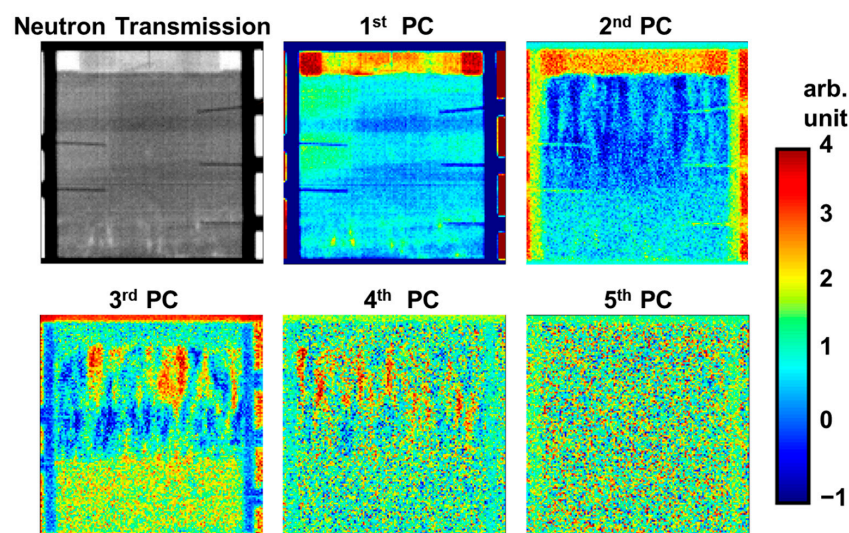


Figure 11. Examples of 100 mm \times 100 mm images of the first through fifth principal components (PC) obtained by the PCA of the neutron transmission spectrum at each pixel (after solidification).

4.4. Solid–Liquid Phase Fraction Imaging of LBE Using Supervised Machine Learning Analysis

Figure 12 shows an example of a single-pixel spectrum used for analysis. It was estimated to be the liquid phase from the thermocouple measurement until 30 min after the solidification process started. The statistical error is large for one-pixel data. Such a spectrum was created pixel by pixel from the TOF transmission image data, and the solid-phase fraction was quantitatively obtained by KNN analysis. Figure 13 shows an example of the result of quantifying the solid-phase fraction, performed on the entire detection area. The linear shadows shown by the blue arrows are thermocouples. To exclude the influence of the background caused by such devices, the solidification process was evaluated in the zone indicated by the red frame. Note that, once the spectral dataset for analysis is prepared, the KNN analysis can be completed in a few minutes, so the effect of time efficiency is extremely high.

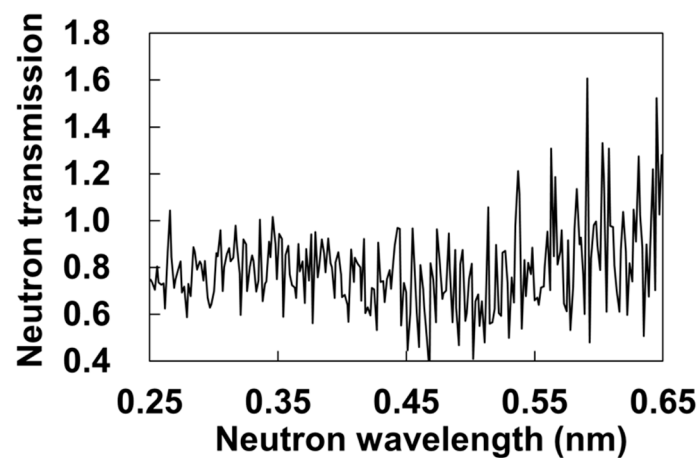


Figure 12. Example of neutron transmission spectrum analyzed by ML analysis. It shows one-pixel data of the liquid-phase zone from the start of the solidification process to 30 min later.

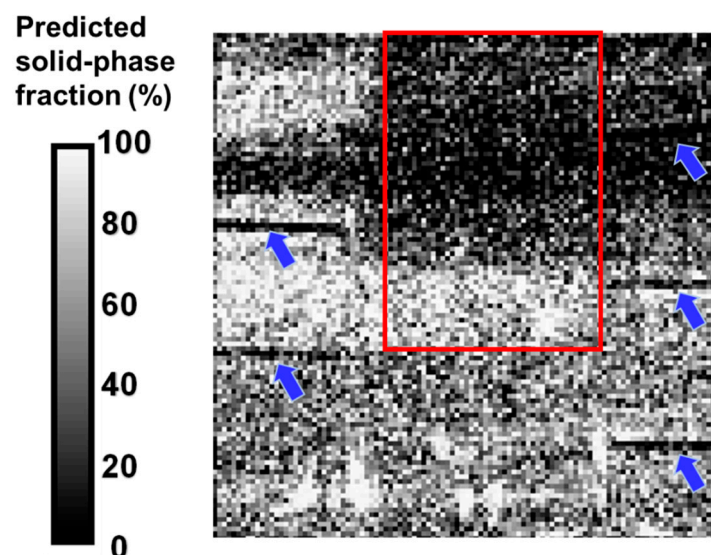


Figure 13. An example of solid–liquid phase fraction imaging ($80\text{ mm} \times 80\text{ mm}$) results obtained by ML analysis (at the beginning of the solidification process). The black shadows at the blue arrow positions in the figure are the thermocouples in the sample. To eliminate the influence of the thermocouples and the insulation material around the sample container, the data inside the red frame were used for the analysis to determine the movement of the solid–liquid interface.

Some examples of the solid-phase fraction map and average in the horizontal direction inside the red-frame zone are shown in Figure 14. In the average diagram on the right in the figure, the horizontal axis is the solid-phase fraction and the vertical axis is the distance from the top of the sample. Initially, the lower zone of the image corresponds to 100% solid phase and the upper zone to 100% liquid phase. In the average diagram, the analyzed results of the solid- and liquid-phase zones do not show perfect 100% and 0% values, but this may be due to the influence of statistical errors and residual background components. However, as shown in the diagram, each zone shows moderately constant values and can be used to determine the solid- and liquid-phase zones. In addition, there is a zone of gradual change in phase fraction between these two zones, which is considered the zone where the solid–liquid interface moved within 30 min. This interfacial zone moves to the top over time, until eventually the whole area becomes solid.

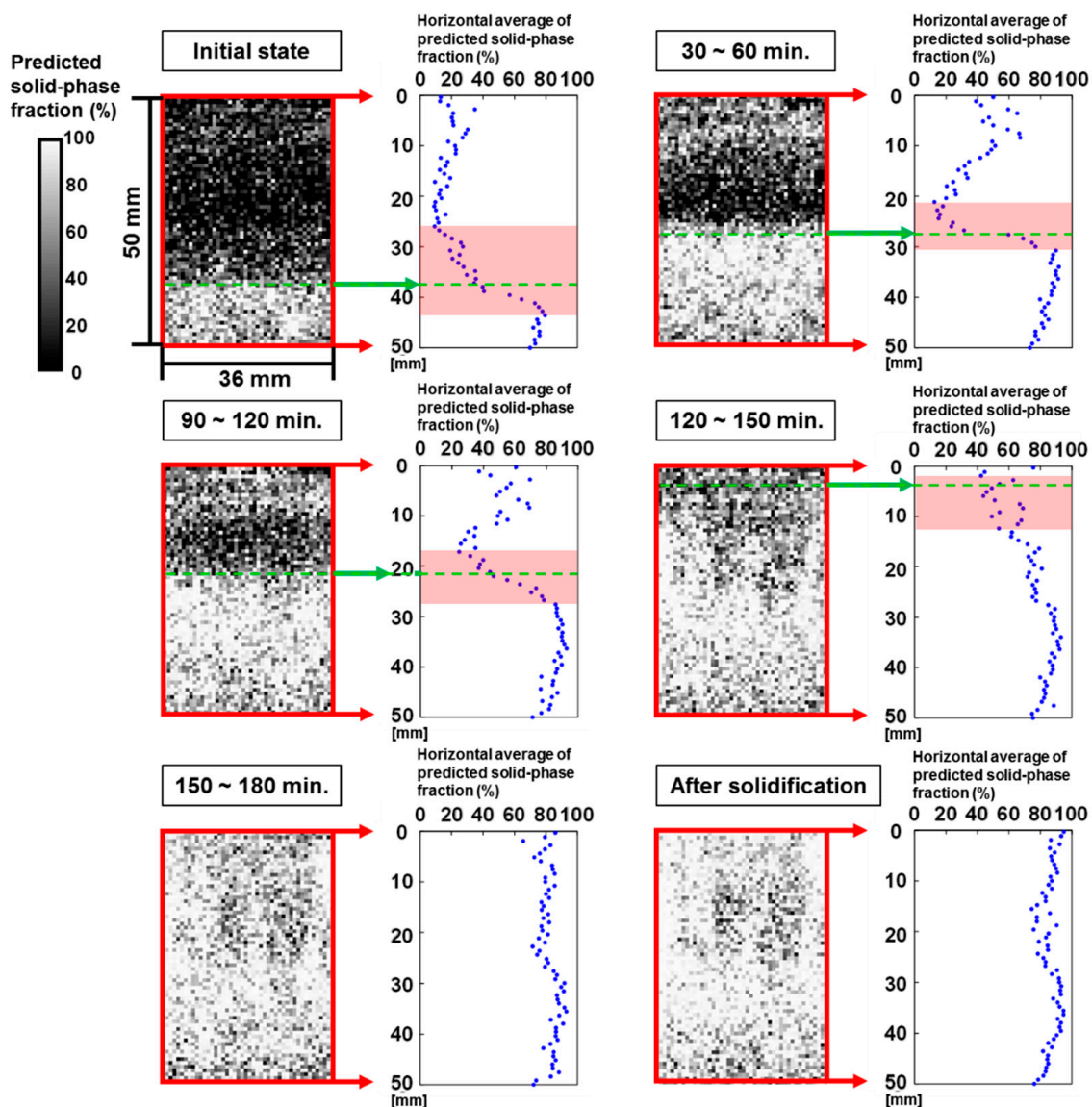


Figure 14. Results of solid-phase fraction imaging during the solidification. For each time period, the left figure shows the mapped real-space distribution of the solid-phase fraction and the right figure shows its horizontal average. On the right, the colored area indicates the location of the solid–liquid coexistence zone, while the green line indicates the interface location estimated by linear fitting where the solid-phase was 50%.

Subsequently, the average solid–liquid interfaces were estimated by the position of the 50% solid-phase fraction for all analyzed images. The results are shown in Figure 15. The horizontal axis shows the time since the solidification process began, and the vertical axis shows the position from the top of the sample. The results are relatively consistent with the time dependence results of a previous study [26,27] estimated from neutron transmission images. Based on the results, we estimated the solidification velocity during the experiment, and this is shown in Figure 16. The result of this study shows a tendency for the solidification velocity to gradually increase with time, which may be reasonable considering the decrease in the liquid-phase volume over time, i.e., the decrease in heat storage capacity.

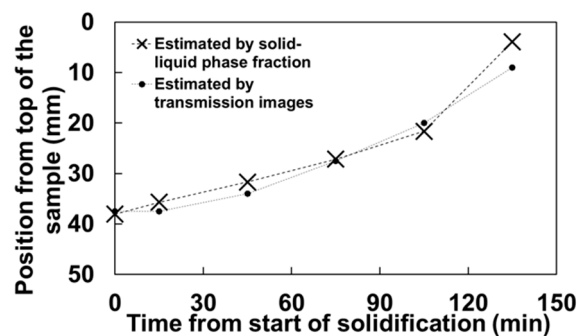


Figure 15. Time dependence of the averaged solid–liquid interface position (50% solid-phase fraction) according to ML analysis. The interface position estimated from the neutron transmission image is also shown as a comparison [26,27].

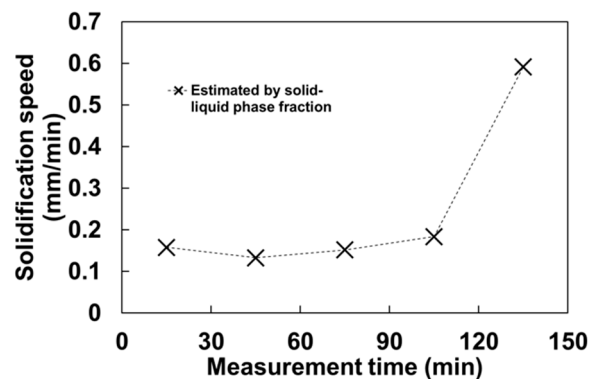


Figure 16. Solidification velocity estimated from the averaged solidification interface position obtained by ML analysis.

5. Conclusions

In this study, the application of ML to neutron transmission spectroscopic imaging (NTSI) was investigated using solid–liquid phase fraction analysis as an example. The target of the analysis was the evaluation of the solid–liquid interface movement during the solidification process of LBE. In supervised ML, creating training data can be an issue, but in this study we assembled the training dataset using actual experimental data with a wide variety of solid–liquid phase fractions and crystalline textures of the solid phase without model calculation.

For the ML analysis, we adopted two steps—the dimensionality reduction of neutron transmission spectrum by PCA and spectrum pattern analysis by KNN—with the goal of reducing the computational cost as much as possible. For the examination of this ML analysis, the evaluating spectrum dataset was created by first setting the phase fraction between the simulated solid spectra and experimental liquid spectrum. Then, the ML was

examined with a dataset that contained Gaussian distributed random errors to simulate statistical errors. As a result, it was found that PCA can evaluate the solid–liquid phase fraction of LBE with principal components up to several dimensions, while KNN can evaluate the fraction with the same accuracy as that found for the training data.

As a result of applying this ML analysis to the actual experimental data, it was possible to evaluate the solid-, liquid-, and mixed-interface zones quantitatively based on the phase fraction. The average position of the solid–liquid interface was generally consistent with the results of a previous study [26,27]. The ML analysis introduced here was for the simple problem of shape recognition about the mixture of the spectra of solid- and liquid-phases, which have clearly different shapes. To extend ML to general spectrum analysis, there are many issues such as accumulation of training data. We are considering applying ML to NTSI spectral analysis by limiting the number of parameters of the crystal texture.

Thus, the application of ML analysis to NTSI seems to have potential. In unsupervised ML, the visualization of PCA results is also expected to be a bridge to detailed crystallographic analysis because it can clearly show the characteristics of the transmission spectrum. Further, the result that the measured spectra can be used as the training data in supervised ML is advantageous to the NTSI, which can collect the number of spectra with multiple pixels on a two-dimensional detector. The effectiveness of the use of experimental spectra for training data, as revealed in this study, shows the potential of ML analysis using large databases of previous experimental data for new experimental data.

Author Contributions: Conceptualization, T.K., K.H. and H.S.; methodology, T.K., K.H., H.S., K.O. and Y.S.(Yuta Suzuki); software, K.H.; validation, T.K. and H.S.; formal analysis, T.K., K.H. and H.S.; investigation, H.S., D.I. and Y.S.(Yasushi Saito); resources, D.I. and Y.S.(Yasushi Saito); data curation, T.K. and K.H.; writing - original draft, T.K.; writing - review & editing, K.H., H.S., K.O., Y.S.(Yuta Suzuki), D.I. and Y.S.(Yasushi Saito); visualization, T.K., K.H. and H.S.; supervision, T.K. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are available on request from the corresponding author.

Acknowledgments: The authors would like to thank T. Shinohara (J-PARC) and the J-PARC MLF BL22 RADEN group members for their assistance with the experiment.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Santisteban, J.R.; Edwards, L.; Fitzpatrick, M.E.; Steuwer, A.; Withers, P.J.; Daymond, M.R.; Johnson, M.W.; Rhodes, N.; Schooneveld, E.M. Strain imaging by Bragg edge neutron transmission. *Nucl. Instrum. Methods. Phys. Res. Sect. A* **2002**, *481*, 765–768. [[CrossRef](#)]
2. Kockelmann, W.; Frei, G.; Lehmann, E.H.; Vontobel, P.; Santisteban, J.R. Energy-selective neutron transmission imaging at a pulsed source. *Nucl. Instrum. Methods. Phys. Res. Sect. A* **2007**, *578*, 421–434. [[CrossRef](#)]
3. Kiyonagi, Y.; Kamiyama, T.; Takada, O.; Morita, K.; Iwase, K.; Suzuki, S.; Furusaka, M.; Hiraga, F.; Tomioka, S.; Mishima, K.; et al. Imaging and texture observation of materials by using a pulsed neutron spectroscopic transmission method. *Nucl. Instrum. Methods. Phys. Res. Sect. A* **2009**, *600*, 167–169. [[CrossRef](#)]
4. Sato, H.; Takada, O.; Iwase, K.; Kamiyama, T.; Kiyonagi, Y. Imaging of a spatial distribution of preferred orientation of crystallites by pulsed neutron Bragg edge transmission. *J. Phys. Conf. Ser.* **2010**, *251*, 012070. [[CrossRef](#)]
5. Sato, H.; Sato, T.; Shiota, Y.; Kamiyama, T.; Tremsin, A.S.; Ohnuma, M.; Kiyonagi, Y. Relation between Vickers Hardness and Bragg-Edge Broadening in Quenched Steel Rods Observed by Pulsed Neutron Transmission Imaging. *Mater. Trans.* **2015**, *56*, 1147–1152. [[CrossRef](#)]
6. Sato, H.; Kamiyama, T.; Kiyonagi, Y. A Rietveld-Type Analysis Code for Pulsed Neutron Bragg-Edge Transmission Imaging and Quantitative Evaluation of Texture and Microstructure of a Welded α -Iron Plate. *Mater. Trans.* **2011**, *52*, 1294–1302. [[CrossRef](#)]
7. Kamiyama, T.; Iwase, K.; Sato, H.; Harjo, S.; Ito, T.; Takata, S.-i.; Aizawa, K.; Kiyonagi, Y. Microstructural Information Mapping of a Plastic-deformed α -iron Plate during Tensile Tests Using Pulsed Neutron Transmission. *Phys. Procedia* **2017**, 50–57. [[CrossRef](#)]

8. Kamiyama, T.; Narita, Y.; Sato, H.; Ohnuma, M.; Kiyanagi, Y. Structural Change of Carbon Anode in a Lithium-ion Battery Product Associated with Charging Process Observed by Neutron Transmission Bragg-edge Imaging. *Phys. Procedia* **2017**, *88*, 27–33. [CrossRef]
9. Butler, K.T.; Davies, D.W.; Cartwright, H.; Isayev, O.; Walsh, A. Machine learning for molecular and materials science. *Nature* **2018**, *559*, 547–555. [CrossRef]
10. Fu, W.; Hopkins, W.S. Applying Machine Learning to Vibrational Spectroscopy. *J. Phys. Chem. A* **2018**, *122*, 167–171. [CrossRef]
11. Suzuki, Y.; Hino, H.; Kotsugi, M.; Ono, K. Automated estimation of materials parameter from X-ray absorption and electron energy-loss spectra with similarity measures. *NPJ Comput. Mater.* **2019**, *5*, 39. [CrossRef]
12. Kamuda, M.; Zhao, J.; Huff, K. A comparison of machine learning methods for automated gamma-ray spectroscopy. *Nucl. Instrum. Methods. Phys. Res. Sect. A* **2020**, *954*, 161385. [CrossRef]
13. Hosseini, S.A. Neutron spectroscopy with soft computing: Development of a computational code based on Support Vector Machine (SVM) for reconstruction of neutron energy spectrum. *J. Instrum.* **2019**, *14*, P02008. [CrossRef]
14. Hey, T.; Butler, K.; Jackson, S.; Thiyagalingam, J. Machine learning and big scientific data. *Phil. Trans. R. Soc. A* **2020**, *378*, 20190054. [CrossRef]
15. Chen, Z.; Andrejevic, N.; Drucker, N.; Nguyen, T.; Xian, R.P.; Smidt, T.; Wang, Y.; Ernstorfer, R.; Tennant, A.; Chan, M.; et al. Machine Learning on Neutron and X-Ray Scattering. *arXiv* **2021**, arXiv:2102.03024. Available online: <https://ui.adsabs.harvard.edu/abs/2021arXiv210203024C> (accessed on 25 June 2021).
16. Song, G.; Porcar, L.; Boehm, M.; Cecillon, F.; Dewhurst, C.; Le Goc, Y.; Locatelli, J.; Mutti, P.; Weber, T. Deep Learning Methods on Neutron Scattering Data. *EPJ Web. Conf.* **2020**, *225*, 01004. [CrossRef]
17. Takada, O.; Kamiyama, T.; Kiyanagi, Y. Study on phase transition of Pb-Bi eutectic alloy by neutron transmission spectroscopy. *J. Nucl. Mater.* **2010**, *398*, 129–131. [CrossRef]
18. Zhang, J. Lead-bismuth eutectic (LBE): A coolant candidate for Gen. IV advanced nuclear reactor concepts. *Adv. Eng. Mater.* **2014**, *16*, 349–356. [CrossRef]
19. Agostini, P.; Baicchi, E.; Zucchini, A.; Benamati, G. The re-crystallization issue in lead-bismuth technology. *J. Nucl. Mater.* **2004**, *335*, 275–279. [CrossRef]
20. Zucchini, A.; Agostini, P.; Baicchi, E. Lead-bismuth eutectic recrystallization studies for the Megapie target. *J. Nucl. Mater.* **2005**, *336*, 291–298. [CrossRef]
21. Khairulin, R.A.; Lyapunov, K.M.; Mozgovoï, A.G.; Stankus, S.V. Crystallization and relaxation phenomena in the bismuth-lead eutectic. *J. Alloy. Compd.* **2005**, *387*, 183–186. [CrossRef]
22. Van der Maaten, L.; Postma, E.; Herik, H. Dimensionality Reduction: A Comparative Review. *J Mach. Learn. Res. JMLR* **2007**, *10*, 13.
23. Wold, S.; Esbensen, K.; Geladi, P. Principal component analysis, Chemometr. *Intell. Lab. Syst.* **1987**, *2*, 37–52. [CrossRef]
24. Saravanan, R.; Sujatha, P. A State of Art Techniques on Machine Learning Algorithms: A Perspective of Supervised Learning Approaches in Data Classification. In Proceedings of the 2018 Second International Conference on Intelligent Computing and Control Systems (ICICCS), Madurai, India, 14–15 June 2018. [CrossRef]
25. Shinohara, T.; Kai, T.; Oikawa, K.; Nakatani, T.; Segawa, M.; Hiroi, K.; Su, Y.; Ooi, M.; Harada, M.; Iikura, H.; Hayashida, H.; et al. The energy-resolved neutron imaging system, RADEN. *Rev. Sci. Inst.* **2020**, *91*, 043302. [CrossRef]
26. Ito, D.; Saito, Y.; Sato, H.; Shinohara, T. Visualization of Solidification Process in Lead-bismuth Eutectic. *Phys. Procedia* **2017**, *88*, 58–63. [CrossRef]
27. Ito, D.; Sato, H.; Saito, Y.; Parker, J.D.; Shinohara, T.; Kai, T. Visualization of phase distribution in lead-bismuth eutectic during one-dimensional solidification process. *J. Vis.* **2019**, *22*, 889–895. [CrossRef]
28. Uno, S.; Uchida, T.; Sekimoto, M.; Murakami, T.; Miyama, K.; Shoji, M.; Nakano, E.; Koike, T.; Morita, K.; Satoh, H.; et al. Two-dimensional Neutron Detector with GEM and its Applications. *Phys. Procedia* **2012**, *26*, 142–152. [CrossRef]