Rapid Estimation of Elemental Composition from X-ray Photoelectron Spectroscopy Survey Data Using a Deep Learning Model

In-Ho Lee
Korea Research Institute of Standards and Science, Daejeon 34113, Korea

Jeong Hoon Choi and Yongsup Park*
Department of Physics, Kyung Hee University, Seoul 02447, Korea
(Dated: November 5, 2023)

The research presents the development and evaluation of a deep learning model designed for the inference of elemental composition from X-ray photoelectron spectroscopy data. To accomplish this, a synthetic dataset containing 78,600 spectra was generated using XPS parameter databases and electron scattering theory with the transport approximation. The synthetic data encompasses substances with varying numbers of atomic species, ranging from 2 to 5. Model performance and uncertainty were assessed through Monte Carlo dropout predictions during testing. Our developed model exhibited remarkable precision in predicting both carbon contamination and composition vector components, culminating in an impressive coefficient of determination ($^2 = 0.999$) that aligns with the test dataset.

I. INTRODUCTION

X-ray photoelectron spectroscopy (XPS) method is widely used in characterizing materials. XPS analysis is the measurement of the composition and atomistic bonding properties of materials. By irradiating a sample with X-rays we obtain the kinetic energy and intensity of photoelectrons emitted by the photoelectric effect[1]. X-ray energy can be absorbed by one of the core electrons. The energy needed to cause the core electron to be emitted and subsequently detected is characteristic of each element. This characteristic feature allows the use of binding energy to identify the elements present on the surface of the material. Peak intensity at a specific binding energy usually displays the total number of photoelectron counts per second. In some cases, overlapping or unresolved peaks may appear on the binding energy axis. It is essential to examine the core levels and Auger lines for each atom.

XPS can detect all elements except hydrogen and helium with detection limits of approximately 0.1%–1%. The relative amounts of the detected elements within the analysis volume can be extracted from the intensities of the photoelectron peaks. Peak positions, widths, relative heights, and other information are necessary to determine the composition of the substance of interest. Since XPS is extremely surface sensitive, care must be taken to avoid surface contamination. Therefore, it is important to observe the changes in XPS data before and after the surface treatment of the material. Since each element has its own unique binding energy, comparing the peaks and binding energies in the spectrum can be used to determine the composition of the elements present on the sample surface. Binding energy is characteristic of the chemical environment of the core-excited atom corresponding structural characterization. Also, when the chemical bonding state of an atom changes, the binding energy usually changes by a few eV, so the chemical bonding state can be inferred from this change. In addition, XPS simultaneously provides chemical information on the chemical structure, degree of carbon contamination, and oxidation state of the sample. Users typically utilize database-driven software that relies on equipment to obtain the elemental proportions of a material. Computational prediction of core-electron binding energies is also developed[2, 3].

XPS results can be challenging to interpret, in general, although there is a database for binding energy values[4, 5]. Sometimes XPS data are often misinterpreted in the literature. Many factors in the experiment affect binding energies in the XPS method. The binding energies of different atomic features can even overlap, further complicating the analysis. In fact, the generation of photoelectrons is closely related to processes that result from X-ray bombardment of a surface including emission of a photoelectron, X-ray fluorescence, and emission of an Auger electron.

By obtaining a variety of experimental measurement data, it is possible to understand the general XPS measurement data. It is also possible to infer XPS spectra of materials with arbitrary compositions from sufficient data. This is known as the traditional XPS analysis and is widely utilized in both materials physics and industrial research. Similarly, research has recently been conducted on inferring composition ratios using deep learning method[6]. It can be seen that the ability to make inferences from data can be achieved through machine learning.

Recent advances in artificial intelligence have made it possible to make systematic inferences that have never been easily attempted before[7]. We started our research with the question: can composition ratio prediction using artificial neural networks reach the level of prediction of XPS expert-level analysis? In this work, we developed

* parky@khu.ac.kr

and tested a model for correlating XPS data with chemical composition using probabilistic inference supported by recent machine learning methods. At this point, it would be natural to try to leverage artificial intelligence to build models for predicting carbon contamination and composition ratio that are as accurate as those of long-time experts.

Section II provides machine learning details. Section III presents and discusses model performance. Finally, Section IV concludes the present work.

II. COMPUTATIONAL METHODS

We first prepared a synthetic dataset, then built a model using a modern deep neural network and optimized the model parameters using both training and validation datasets. Finally, we used the model to show the precision of our predictions on a test dataset that was not included in the training and validation. The National Institute of Standards and Technology (NIST) provides several types of XPS-related data and a useful program for the simulation of XPS spectra (Simulation of Electron Spectra for Surface Analysis (SESSA))[8, 9]. It is possible to generate a synthetic dataset for XPS by taking into account composition and carbon contamination[10]. The neural network we built is faithful to modern deep learning techniques. It was trained on a synthetic dataset and computed on a validation dataset, so there was no overfitting.

A Monte Carlo dropout approach[11] was utilized to predict carbon contamination levels and composition ratios. Typically, a Monte Carlo dropout approach is utilized for more reliable predictions. Monte Carlo dropout is a technique that extends the use of dropout to the inference mode of a neural network. Unlike the usual case, we also take advantage of the stochastic dropout feature in prediction mode. It is solely based on a different set of neurons that are dropped out. In addition, for each prediction, we have both mean and variance. It utilizes the dropout layer, which is also enabled in inference, to make multiple predictions on the same input. From these multiple predictions, the mean and variance can be calculated. To get a single mean and variance, we made 200 independent predictions. The averaging step smooths out the noise introduced by the dropout layers and can lead to more accurate predictions.

A. synthetic data

In order to train a deep learning model with complex and many parameters, a large amount of high-quality data is required. It was impossible to make repeated measurements on such a large number of samples through real experiments, so we decided to use synthetic data that exquisitely mimics the characteristics of real XPS. We generated a synthetic dataset made of $75000 + 600 + 3000$ spectra based on XPS parameter databases and electron scattering theory within the transport approximation[10]. We tried to express characteristics similar to data that can be obtained through actual experiments by adding the SESSA database and some realistic features. The SESSA provides synthetic data using element parameters as arguments. We consider random possible combinations of elements from the list of Li, Be, B, C, ..., and Bi atoms except Pm, which is not supported. Each virtual material is composed of a random number from 2 to 5 of atomic species, with variable stoichiometry ratios. Only Al k_α source is used in the present simulations.

Even if great care is taken during the actual experiment, contamination from airborne substances such as Carbon and Oxygen is inevitable. In the experiment, the carbon contamination leads to an overall lower XPS intensity[12]. We represent carbon contamination on data by stacking the separated uniformly random depth (from 0 to 40 Å) C_5O layer flatly on top of the elements layer. The chemical shift is related to the electronegativity of the nearest neighbor atoms.[13] This chemical shift occurs in each peak in all element combinations, but it is difficult to predict for all substances. We expressed the chemical shift phenomenon in the data by applying a random shift of up to 10eV to each peak. For the carbon and oxygen peak, a shift of up to 0.5 eV was applied. Also, there is a significant difference in the contribution of each element in XPS. Therefore, if the ratio of elements is used as is, it is impossible to calculate fairly for each element in the loss calculation for the composition ratio. Therefore, we created all spectra of the elements used through SESSA and used the intensity of the maximum peak of the element to create new relative intensity data T_i and use it. This relative intensity is used to redefine labels for the Deep Neural Network model. We use label define as $\bar{y}_i = \bar{q}_i T_i (\sum_i \bar{q}_i T_i)^{-1}$.[14]

The completed spectra kinetic energy range was $400 - 1486$ eV on a 2048 energy point grid. The present network takes as input the 2048 spectral points ($\{x_j\}, j = 1, 2, 3, ..., 2048.$) intensity and produces two outputs, the normalized carbon contamination level $c (0 \leq c \leq 1)$, indicating a normalized thickness ($c = 1$ means 40Å.), and the normalized composition ratios ($\{y_e\}, 0 \leq y_e \leq 1, \sum_{e=1}^{81} y_e = 1, e = 1, 2, 3, ..., 81.$) of the 81 element. The energy grid we used is generic without loss of generality. This is because even if counting data is obtained with different energy intervals, the XPS information can be converted to the 2048 energy grid with appropriate transformations.

SESSA offers both a Command Line Interface(CLI) and a Graphical User Interface(GUI). Using the CLI, we could obtain XPS data for two-layer materials for iteratively determined parameters(elemental composition, contaminant layer thickness). The entire data generation process was executed on a 12-core processor. It involved three processes for two data processing and one process for management. The data processing processes

have been separated depending on whether they depend on SESSA or not. Data generation took around $10-20$ seconds per sample on average which highly correlated to SESSA, and creating 75k samples took approximately 10 days.

B. deep learning model

The single neural network we used predicts each of the two by itself. The first is the degree of carbon contamination($0 \leq c \leq 1$) and the second is the composition ratio($0 \leq y_e \leq 1$, $\sum_{e=1}^{81} y_e = 1$, $e = 1, 2, 3, \ldots, 81$). As explained earlier, H and He are excluded from the study. We built a model that simultaneously optimizes the regression loss functions (contamination degree and composition ratio) using multiple one-dimensional convolutional neural network layers, pooling layers, multi-head attention layers, dense layers, and dropout layers. Our model is based on TensorFlow.

The model was trained using dropout with a dropout rate of 0.1. A 'softmax' activation function for the final layer is used to normalize the outputs so that $\sum_{e=1}^{81} y_e = 1$. A 'sigmoid' activation, is used to identify the level of carbon contamination $0 \leq c \leq 1$. Total number of trainable parameters is 8775775. An 'adam' optimizer[15] and the tuned loss functions allowed for robust training, without overfitting. Using dropout, we can reduce concerted learning among neurons, which may have led to overfitting. Overfitting occurs when the model has a high variance. In this case, the model performs well on the training dataset but does not perform properly on the validation dataset.

The adoption of 'ReLU (rectified linear unit)' may easily be considered one of the few milestones in the deep learning history. Activation function with 'ReLu' permits the routine development of very deep neural networks. We used the 'ReLU' function as the activation function unless it is the last layer that predicts the function value. The last layer for a prediction consists of a dense layer. In general, a good method to prevent overfitting is to use a more complete training dataset. The dataset should cover the entire range of inputs that the model is expected to process. An additional dataset may only be useful if it covers new cases.

We use the Monte Carlo dropout method[11] to make predictions, so enabling probabilistic dropout also works in prediction mode. From these predictions, we get both the predicted value and the uncertainty about the prediction. In the Monte Carlo dropout, multiple predictions can be samples from a probabilistic distribution. Thus, Monte Carlo dropout is a Bayesian interpretation of dropout. Basically, the Monte Carlo dropout method provides a way to obtain uncertainty estimates and model confidence, which can be crucial in many machine learning applications.

C. loss functions

Mean squared logarithmic error(MSLE, L) is considered to be an improvement over using percentage-based errors for training because its numerical properties are better.

$$L = \frac{1}{81} \sum_{e=1}^{81} \left\{ log(1 + y_e) - log(1 + y_e^t) \right\}^2, \qquad (1)$$

where $\{y_e\}$ and $\{y_e^t\}$ are the network outputs and the target values, respectively. It is less sensitive to outliers than mean squared error since the logarithmic transformation compresses the error values. The loss function, L is less sensitive to scale variation as it is a difference of two log values which is the same as a log of the ratio of the values. Due to the loss being log it penalizes underestimates more than overestimates.

In statistics, the coefficient of determination denoted R^2, is the proportion of the variation in the dependent variable (chemical composition) that is predictable from the independent variables (XPS spectrum). In regression, the R^2 coefficient of determination is a statistical measure of how well the regression predictions approximate the real data points. In the best case, the predicted values exactly match the true values, which results in $R^2 = 1$.

$$R^2 = 1 - \frac{\sum_{i=1}^{M} (z_i - z_i^t)^2}{\sum_{i=1}^{M} (z_i - \bar{z^t})^2}, \qquad (2)$$

where $\bar{z^t}$ denotes the mean of the true values, z_i denotes data point, z_i^t denotes true value, and M denotes the number of data points associated with prediction. Even if we define a single predictor that incorporates the carbon contamination (c) and composition vector($\{y_e\}$), we can still calculate a statistically defined R^2 value for many test samples. We can interpret this problem as a regression analysis that predicts 82 values, $c, y_1, y_2, \ldots, y_{81}$. For example, once we have a test dataset, consisting of 1000 examples, then $M = 82 \times 1000$. Our outputs of the present neural network, both contamination (c) and composition ratio ($\{y_e\}$) can be viewed as a regression problem. By utilizing R^2 measures that are widely used in regression analysis, we can quantify how accurately the developed model predicts contamination and composition ratio. Usually, the larger the R^2, the better the regression model fits observations.

D. training dataset, validation dataset, and test dataset

The validation dataset is used during the training mode of the model to provide an unbiased evaluation of the model's performance. The test dataset, on the other hand, is used after the model has been fully trained to assess the model's performance on a completely unseen

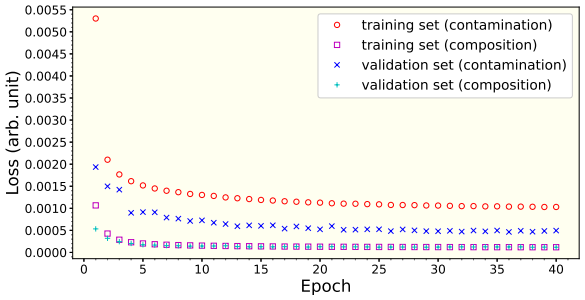

FIG. 1. Loss functions are related to the prediction of the degree of carbon contamination and the prediction of the composition ratio, respectively. Mean squared logarithmic error (MSLE) is used for both contamination prediction and composition prediction. Total number of trainable parameters is 8775775. The training dataset size and the validation dataset size are 75000 and 600, respectively.

dataset. We plotted the loss functions over the training epoch in the Fig.1. We also plotted the loss functions for the validation dataset on the model at the same time. The loss functions are related to the prediction of the degree of carbon contamination and the prediction of the composition ratio, respectively. In our case, with a sufficiently large data size, we verified that the present model does not cause overfitting problems.

Thinking of the prediction of carbon contamination and composition vector components as a regression problem, for our test dataset, we were able to achieve a precision equivalent to $R^2 = 0.999$. Thus, the predicted values are well match the true values for the test dataset. Especially, this precision means that the model is very good at selectively predicting the atomic species that make up the sample. This interpretation is just one analysis of how well the model can differentiate between different atomic species from the XPS data. And so, it is worth performing a maximum absolute error (L_∞-norm) analysis on the predictions, both carbon contamination and composition. The mean absolute error (mean of $\{\mid c - c^t \mid\}$, $0 \leq c \leq 1$) in the carbon contamination prediction for the test dataset is 0.1293. Here, braces represent samples. c and c^t are predicted carbon contamination values and true carbon contamination values, respectively. Here, the test dataset size is 1000. For the same test dataset, the mean of the maximum absolute error (mean of $\{max_e \mid y_e - y_e^t \mid\}$, L_∞-norm) for the composition ratio prediction is 0.0491. Thus, we have a set of results, (0.1293, 0.0491) for error estimations, [(mean of $\{\mid c - c^t \mid\}$, mean of $\{max_e \mid y_e - y_e^t \mid\}$)]. We obtained two sets of similar results, (0.1301, 0.0479) and (0.1423, 0.0531), for two other independent test datasets, respectively.

III. RESULTS AND DISCUSSION

In general, XPS measurement results are dominated by near-surface information, so predicting the extent of surface contamination is a critical factor in sample preparation and interpretation. In our study, the XPS data points were represented as a vector. This is the input to the model and the carbon contamination and composition ratios are the outputs of the model. Instead of the artificial neural network responsible for the prediction being a single function, we can consider probabilistic models. In this case, a single model can provide a probabilistic distribution. From this approach, we can determine the probabilistic distribution of the model's predictions. Of course, this allows us to make more reliable predictions.

Fig.2 (a) shows the XPS data used as an input. The assumption used here is that only XPS data is used for predictions of carbon contamination and composition ratios. We did not use information about how many different atomic species it is composed of. It has the following carbon contamination level and composition ratios. True values for this test case are $c^t = 0.758, (y^t_{\mathrm{Al}}, y^t_{\mathrm{Cl}}, y^t_{\mathrm{Ge}}, y^t_{\mathrm{Nd}}, y^t_{\mathrm{Pb}}) = (0.011, 0.070, 0.348, 0.540, 0.029)$. Fig.2 (b) and (c) show the results predicted by the neural network. In Fig.2 (b), the actual and predicted values of the composition ratios are shown simultaneously for direct comparison. Fig.2 (c) shows the predicted value of the composition ratios (mean values) and the error (2σ, σ: standard deviation) at the same time. Predicted values are $c = 0.780, (y_{\mathrm{Al}}, y_{\mathrm{Cl}}, y_{\mathrm{Ge}}, y_{\mathrm{Nd}}, y_{\mathrm{Pb}}) = (0.010, 0.030, 0.310, 0.552, 0.004)$. For the carbon contamination prediction, we have 2.90% error. If we limit our analysis to the case of predicting only 81 values, we have an $R^2 = 0.990$ value for predicting relative elemental abundances. Thus, compared to the average prediction ability of our model, this case is a fairly large error in the composition prediction. In our calculations, GPU-enabled computations can realistically yield significant acceleration in model building and utilization.

Fig.3 (a) shows the XPS data used as an input. True values for this test case are $c^t = 0.678, (y^t_{\mathrm{Nd}}, y^t_{\mathrm{Tl}}) = (0.475, 0.524)$. In Fig.3 (b), the actual and predicted values of the composition ratios are shown simultaneously for direct comparison. Fig.3 (c) shows the predicted value of the composition ratio (mean values) and the error (2σ) at the same time. Predicted values are $c = 0.682, (y_{\mathrm{Nd}}, y_{\mathrm{Tl}}) = (0.482, 0.507)$. For the carbon contamination prediction, we have 0.58% error. If we limit our analysis to the case of predicting only 81 values, we have an $R^2 = 0.992$ value for predicting relative elemental abundances.

Fig.4 (a) shows the XPS data used as an input. True values for this test case are $c^t = 0.365, (y^t_{\mathrm{N}}, y^t_{\mathrm{Rh}}, y^t_{\mathrm{Ag}}, y^t_{\mathrm{Ce}}) = (0.003, 0.496, 0.258, 0.177)$. In Fig.4 (b), the actual and predicted values of the composition ratios are shown simultaneously for di-

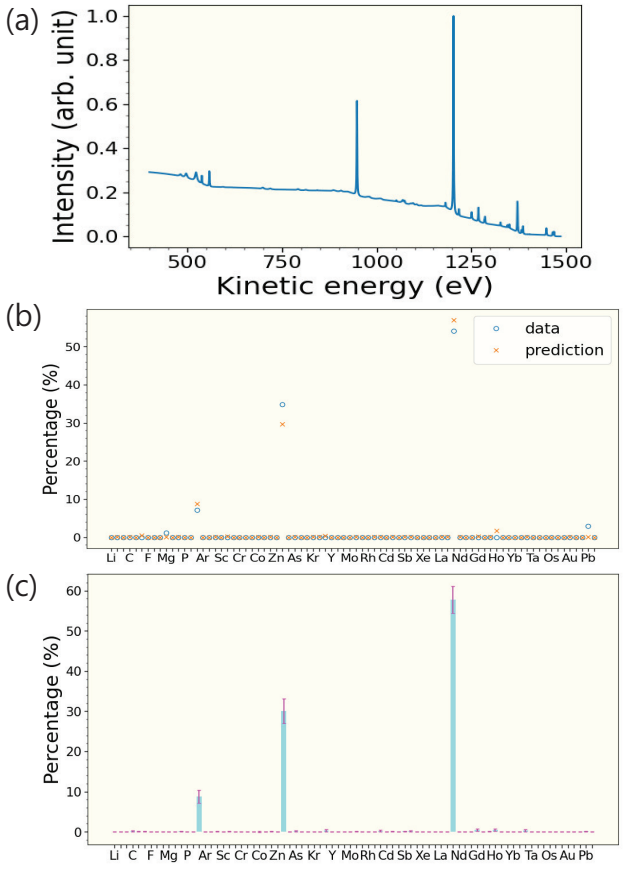

FIG. 2. (a) A typical simulated XPS spectrum is a plot of the number of electrons detected at a specific binding energy. XPS peaks correspond to the electron configuration of the electrons within the atoms. The number of detected electrons in each XPS peak is closely related to the amount of element within the sampling volume. (b) An example of predicting composition using only XPS data is shown. The circle and cross represent the true value and predicted value, respectively. (c) The uncertainty of the prediction can be obtained with the Monte Carlo dropout method. Many different elements are predicted to contribute with very small percentages, but the uncertainty of the prediction must be taken into account.

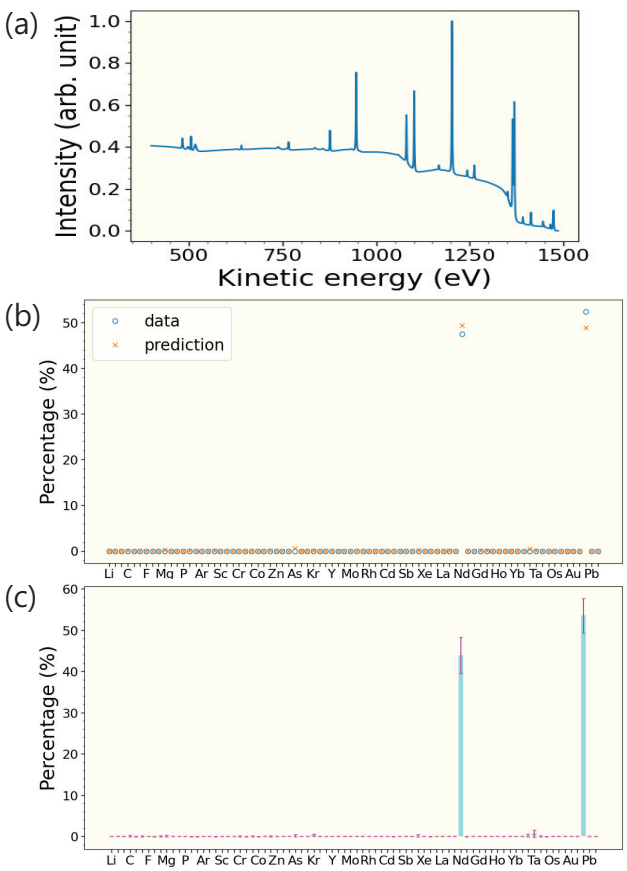

FIG. 3. (a) A typical simulated XPS spectrum is a plot of the number of electrons detected at a specific binding energy. (b) An example of predicting composition using only XPS data is shown. The circle and cross represent the true value and predicted value, respectively. (c) The uncertainty of the prediction can be obtained with the Monte Carlo dropout method.

rect comparison. Fig.4 (c) shows the predicted value of the composition ratio (mean values) and the error (2σ) at the same time. Predicted values are $c = 0.363, (y^t_{\mathrm{N}}, y^t_{\mathrm{Rh}}, y^t_{\mathrm{Ag}}, y^t_{\mathrm{Ce}}) = (0.014, 0.389, 0.328, 0.217)$. For the carbon contamination prediction, we have 0.55% error. If we limit our analysis to the case of predicting only 81 values, we have an $R^2 = 0.982$ value for predicting relative elemental abundances.

An expert in the field who has been working with XPS data for a long time will be able to use existing information to roughly predict the composition and carbon contamination of a sample, respectively. Similarly, our artificial neural network model was trained on XPS data using a 75000 virtual training dataset. The predictive power of this training was validated by training an artificial neural network. After training the model, a single prediction using a single piece of XPS data takes about 7 seconds of CPU computation time (actually, 200 independent predictions) on a typical personal computer. This reduction in composition ratio and carbon contamination prediction time is the basis for XPS measurements to become a new form of material surface analysis in the near future. We found the model we developed to be efficient and effective in predicting elemental composition and contamination without overfitting. With the deep learning techniques we used, many more applications related to XPS measurements will be possible in the future. It depends, but sometimes researchers need very fast analyses. This is because surface contamination is associated with some form of time delay. Therefore, the quick analysis of contamination using deep learning that we are looking for can be very meaningful in some situ-

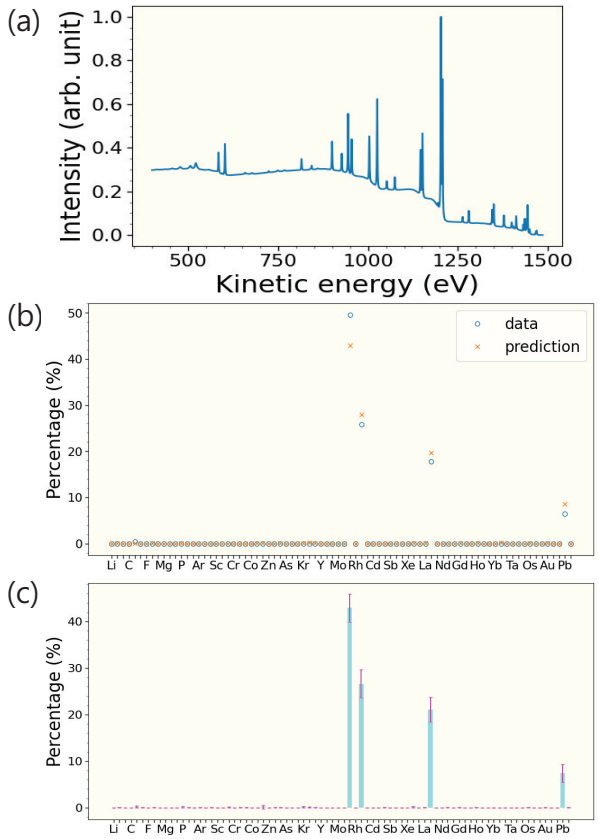

FIG. 4. (a) A typical simulated XPS spectrum is a plot of the number of electrons detected at a specific binding energy. (b) An example of predicting composition using only XPS data is shown. The circle and cross represent the true value and predicted value, respectively. (c) The uncertainty of the prediction can be obtained with Monte Carlo dropout method.

ations.

IV. CONCLUSIONS

To summarise, an accurate and efficient deep learning model was developed to deduce elemental composition from XPS data. It has been demonstrated that the model possesses expert-level predictive capabilities and can be used to infer composition and carbon contamination as a basic application. We produced a synthetic dataset of X-ray photoelectron spectroscopy comprising 78600 spectra, using XPS parameter databases and electron scattering theory within the transport approximation. We considered 2, 3, 4, and 5 as the potential number of atomic species present in a substance. We employed the latest techniques in artificial intelligence to construct a model capable of forecasting the extent of carbon contamination and composition. To predict carbon contamination levels and composition ratios, we utilized a Monte Carlo dropout approach. For every prediction, we obtained both the mean and variance. By employing this approach, we were able to offer an accurate forecast and a range of confidence for the forecast. The model that we devised displays a strong accuracy with respect to the coefficient of determination ($R^2 = 0.999$) for the test dataset when predicting the carbon contamination and composition vector components. The model showed a level of accuracy that could actually be used to distinguish the main elements contained in the actual experimental data but showed significant differences in quantitative ratios. By using the narrow scan data used in actual quantitative analysis, it is expected that a model with high accuracy that can replace the current process can be developed.

This research was supported by the Enhancement of Measurement Standards and Technologies in Physics funded by Korea Research Institute of Standards and Science (No. KRISS-2021-GP2021-0002). This research was also supported by the Commercialization Promotion Agency for R&D Outcomes (COMPA) funded by the Ministry of Science and ICT (MSIT, 1711198537).

[1] C. S. Fadley, X-ray photoelectron spectroscopy: Progress and perspectives, Journal of Electron Spectroscopy and Related Phenomena **178**, 2 (2010).
[2] D. Golze, M. Hirvensalo, P. Hernández-León, A. Aarva, J. Etula, T. Susi, P. Rinke, T. Laurila, and M. A. Caro, Accurate computational prediction of core-electron binding energies in carbon-based materials: A machine-learning model combining density-functional theory and gw, Chemistry of Materials **34**, 6240 (2022).
[3] Q. Sun, Y. Xiang, Y. Liu, L. Xu, T. Leng, Y. Ye, A. Fortunelli, W. A. Goddard III, and T. Cheng, Machine learning predicts the x-ray photoelectron spectroscopy of the solid electrolyte interface of lithium metal battery, The Journal of Physical Chemistry Letters **13**, 8047 (2022).
[4] J. Chastain and R. C. King Jr, Handbook of x-ray photoelectron spectroscopy, Perkin-Elmer Corporation **40**, 221 (1992).
[5] B. V. Crist, Xps in industry—problems with binding energies in journals and binding energy databases, Journal of Electron Spectroscopy and Related Phenomena **231**, 75 (2019).
[6] G. Drera, C. M. Kropf, and L. Sangaletti, Deep neural network for x-ray photoelectron spectroscopy data analysis, Machine Learning: Science and Technology **1**, 015008 (2020).
[7] I. Goodfellow, Y. Bengio, and A. Courville, *Deep learning* (MIT press, 2016).
[8] W. Smekal, W. S. Werner, and C. J. Powell, Simulation of electron spectra for surface analysis (sessa): a novel software tool for quantitative auger-electron spectroscopy and x-ray photoelectron spectroscopy, Surface and Interface Analysis: An International Journal devoted to the

development and application of techniques for the analysis of surfaces, interfaces and thin films **37**, 1059 (2005).

[9] W. S. Werner, W. Smekal, and C. J. Powell, Nist database for the simulation of electron spectra for surface analysis (sessa), in *Version 1.3, Standard Reference Program Database 100, US Department of Commerce* (National Institute of Standards and Technology Gaithersburg, MD, 2011).

[10] W. S. Werner, Electron transport in solids for quantitative surface analysis, Surface and Interface Analysis: An International Journal devoted to the development and application of techniques for the analysis of surfaces, interfaces and thin films **31**, 141 (2001).

[11] Y. Gal and Z. Ghahramani, Dropout as a bayesian approximation: Representing model uncertainty in deep learning, in *international conference on machine learning* (PMLR, 2016) pp. 1050–1059.

[12] S. Evans, Correction for the effects of adventitious carbon overlayers in quantitative xps analysis, Surface and Interface Analysis: An International Journal devoted to the development and application of techniques for the analysis of surfaces, interfaces and thin films **25**, 924 (1997).

[13] K. Siegbahn, C. Nordling, G. Johansson, J. Hedman, P. Hedén, K. Hamrin, U. Gelius, T. Bergmark, L. Werme, R. Manne, *et al.*, *ESCA: Applied to Free Molecules* (North-Holland Publishing Company, 1969).

[14] G. Drera, C. M. Kropf, and L. Sangaletti, Deep neural network for x-ray photoelectron spectroscopy data analysis, Machine Learning: Science and Technology **1**, 015008 (2020).

[15] D. P. Kingma and J. Ba, Adam: A method for stochastic optimization, arXiv preprint arXiv:1412.6980 (2014).