

A Survey on Machine Learning in Chemical Spectral Analysis

Dongfang Yu and Jinwei Wang*

School of Computer and Software, Nanjing University of Information Science and Technology, Nanjing, 210044, China

*Corresponding Author: Jinwei Wang. Email: wjwei_2004@163.com

Received: 15 August 2020; Accepted: 21 September 2020

Abstract: Chemical spectral analysis is contemporarily undergoing a revolution and drawing much attention of scientists owing to machine learning algorithms, in particular convolutional networks. Hence, this paper outlines the major machine learning and especially deep learning methods contributed to interpret chemical images, and overviews the current application, development and breakthrough in different spectral characterization. Brief categorization of reviewed literatures is provided for studies per application apparatus: X-Ray spectra, UV-Vis-IR spectra, Micro-scope, Raman spectra, Photoluminescence spectrum. End with the overview of existing circumstances in this research area, we provide unique insight and promising directions for the chemical imaging field to fully couple machine learning subsequently.

Keywords: Machine learning; chemical imaging; convolutional neural networks

1 Introduction

The strong ability of machine learning [1] to learn high-level features motivates scientists to use this method in all walks of life. This is a direct consequence of the recent breakthroughs resulting from its application across a wide variety of scientific fields, such as data mining [2], computer vision [3], natural language processing [4], biometric recognition [5], medical diagnosis [6], brain circuits studies [7], particle physics [8], securities market analysis, DNA analysis [9], voice and handwriting recognition [10,11], etc. Recently it has also grabbed an increasing interest of scientists in chemical analysis [12]. Machine learning holds out promises for in the field of analytic chemistry. The machine learning architecture—the core of artificial intelligence—enables machines to realize complicated mathematical models of to acquire accurate information of chemical spectral. The input image data are performed by non-linear and/or linear functions of hierarchically models. Treating these functions as data processing ‘layers’, the hierarchical use of a larger number of such layers also inspires the name ‘deep’ learning. The common goal of deep learning methods is to iteratively learn the parameters of the computational model using a training data set such that the model gradually gets better in performing a desired task, e.g., classification; over that data under a specified metric. The computational model itself generally takes the form of an artificial neural network (ANN) [13] that consists of multiple layers of neurons/perceptrons.

Deep convolutional networks have now become the technique of choice. Once trained for a particular task, the deep learning models are also able to perform the same task accurately using a variety of previously unseen data (i.e., testing data). This strong generalization ability of deep learning currently makes it stand out of the other machine learning techniques. Learning of the parameters of a deep model is carried out with the help of back-propagation strategy [14] that enables some form of the popular gradient descent technique [15,16] to iteratively arrive at the desired parameter values. Updating the model parameters using the complete training data once is known as a single epoch of network/model training. contemporary deep learning models are normally trained for hundreds of epochs before they can be deployed.



A crucial step in the design of such systems is the extraction of discriminant features from the images. This process is still done by human researchers and, as such, one speaks of systems with handcrafted features. A logical next step is to let computers learn the features that optimally represent the data for the problem at hand. This concept lies at the basis of many deep learning algorithms: models (networks) composed of many layers that transform input data (e.g., images) to outputs (e.g., defects present/absent) while learning increasingly higher-level features. The most successful type of models for image analysis to date are convolutional neural networks (CNNs). CNNs contain many layers that transform their input with convolution filters of a small extent. The first successful real-world application in LeNet for handwritten digit recognition. Despite these initial successes, the use of CNNs did not gather momentum until various new techniques were developed for efficiently training deep networks. The watershed was the contribution of Krizhevsky et al. to the ImageNet challenge in December 2012 [17]. The proposed CNN, called AlexNet, won that competition by a large margin. In subsequent years, further progress has been made using related but deeper architectures [18].

In this paper, we provide a comprehensive review of the recent machine learning techniques in chemical imaging analysis, focusing mainly on the common spectral characterization. We categorize these techniques under different characterization instruments. Analyzing the reviewed literature, we establish 'lack of appropriately annotated largescale datasets for chemical imaging tasks as the fundamental challenge (among other challenges) in fully exploiting deep learning for those tasks. We then provide guidelines to deal with this and other challenges in chemical image analysis using machine learning. This review also touches upon the available public datasets to train machine learning models for the chemical imaging tasks. Considering the lack of in-depth comprehension of machine learning framework by the broader chemical community, this article also provides the understanding of the core technical concepts related to an appropriate level.

2 Outline of Machine Learning Methods

The purpose of this section is to present the general types of machine learning techniques: supervised and unsupervised learning algorithms. Furthermore, deep learning methods are mainly introduced. Supervised learning algorithms provide the train dataset $D = \{x, y\}_{n=1}^N$ of training example x and its label y . In that case, the training example x belong to different class of data. Label y is a vector with continuous values. Actually, the goal of supervised learning is to find a computational model θ that correctly predict the label y based on the loss function $L(y, \hat{y})$.

In comparison with supervised learning, unsupervised one without labels is used to find patterns. The aim of unsupervised learning is to cluster the data samples into different group based on the similarities of their intrinsic characteristics-e.g. clustering pixel of color images based on their RGB values. Similar to the supervised learning tasks, models for unsupervised learning task can also take advantage of minimizing a loss function. In the context of deep learning, this loss function is normally designed such that the model learns an accurate mapping of an input signal to itself. Once the mapping is learned, the model is used to compute compact representations of data samples that cluster well. Deep learning framework has also been found very effective for unsupervised learning.

Along with supervised and unsupervised learning, other machine learning types include semi-supervised and reinforcement learning. Informally, semi-supervised learning computes models using the training data that provides label only for its smaller subsets. On the other hand, reinforcement learning provides a kind of supervision for the learning problem in terms of rewards or punishments to the algorithm. Due to their remote relevance to the task in chemical spectrum, we do not provide further discussion on these categories.

2.1 Artificial Neural Networks

An artificial neural network (ANN) is a hierarchical composition of basic computational elements known as neurons. Neural networks are a type of learning algorithm which forms the basis of most deep learning methods. A neural network comprises of neurons or units with some activation a and parameters

$\Theta = \{W, B\}$, where W is a set of weights and B a set of biases. The activation represents a linear combination of the input x to the neuron and the parameters, followed by an element-wise nonlinearity $\sigma(\cdot)$, referred to as a transfer function:

$$a = \sigma(w^T x + b). \quad (1)$$

Typical transfer functions for traditional neural networks are the sigmoid and hyperbolic tangent function. The multi-layered perceptrons (MLP), the most well-known of the traditional neural networks, have several layers of these transformations:

$$f(x, \Theta) = \sigma(W^T \sigma(W^T \dots \sigma(W^T x + b)) + b). \quad (2)$$

Here, W is a matrix comprising of columns w_k , associated with activation k in the output. Layers in between the input and output are often referred to as ‘hidden’ layers. When a neural network contains multiple hidden layers, it is typically considered a ‘deep’ neural network, hence the term ‘deep learning’.

At the final layer of the network the activations are mapped to a distribution over classes $P(y|x; \Theta)$ through:

$$P(y|x; \Theta) = \text{softmax}(x; \Theta) = \frac{e^{w_i^T x + b_i}}{\sum_{k=1}^K e^{w_k^T x + b_k}} \quad (3)$$

where w_k indicates the weight vector leading to the output node associated with class i . A schematic representation of three-layer MLP is shown in Figure 1.

Maximum likelihood with stochastic gradient descent is currently the most popular method to fit parameters Θ to a dataset D . In stochastic gradient descent a small subset of the data, a mini-batch, is used for each gradient update instead of the full data set. Optimizing maximum likelihood in practice amounts to minimizing the negative log-likelihood:

$$\arg \min_{\Theta} - \sum_{n=1}^N \log[P(y_n|x_n; \Theta)]. \quad (4)$$

This results in the binary cross-entropy loss for two class problems and the categorical cross-entropy for multi-class tasks. A downside of this approach is that it typically does not optimize the quantity we are interested in directly, such as area under the receiver operating characteristic curve or common evaluation measures for segmentation, such as the dice coefficient.

2.2 Convolutional Neural Networks

In the context of deep learning techniques for image analysis, CNNs [17,19] are of the primary importance. Similar to the standard ANNs, CNNs consist of multiple layers. However, instead of simple perceptron layers, we encounter three different kinds of layers in these networks (a) convolutional layers, (b) pooling layers, and (c) fully connected layers. We describe these layers below, focusing mainly on the Convolutional layers that are the main source of strength for CNNs as shown in Fig. 1.

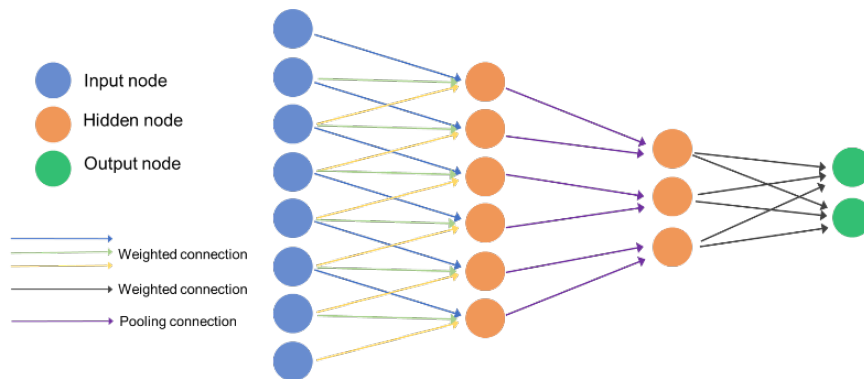


Figure 1: Schematic diagram of convolutional neural network

Convolutional layers: The aim of convolution layers is to learn weights of the so-called convolutional kernels/filters that can perform convolutional operations on images. Traditional images analysis has a long history of using such filter to highlight/extract different image features, such as Sobel filter for detecting edges in images. However, before CNNs, these filters needed to be designed by manually setting the weights of the kernel in a careful manner. The breakthrough that CNNs provided is in the automatic learning of these weights under the neural network setting. In 2D setting (e.g., grey-scale images), this operation involves moving a small window (i.e., kernel) over a 2D grid (i.e., image). In each moving step, the corresponding elements of two grids get multiplied and summed up to compute a scalar value. Concluding the operation results in another 2D-grid, referred as the feature/activation map in the CNN literature. In 3D setting, the sample steps are performed for the individual pairs of the corresponding channels of the 3D volumes, and the resulting feature maps are simply added to compute a 2D map as the final output.

Pooling layers: The main objective of a pooling layer is to reduce the width and height of the activation maps in CNNs. The basic concept is to compute a single output value v for a small $\eta_p \times \eta_p$ grid in the activation map, where v is simply the maximum or average value of that grid in the activation map. Based on the used operation, this layer is often referred as max-pooling or average-pooling layer. Interestingly, there are no learnable parameters associated with a pooling layer. Hence, this layer is sometimes seen as a part of convolutional layer. For instance, the popular VGG-16 network does not see pooling layer as a separate layer, hence the name VGG-16. On the other hand, other works, that use VGG-16 often count more than 16 layers in this network by treating the pooling layer as a regular network layer.

Fully connected layers: These layers are the same as the perceptron layers encountered in the standard ANNs. The use of multiple convolutional and pooling layers in CNNs gradually reduces the size of resulting activation maps. Finally, the activation maps from a deeper layer are rearranged into a vector which is then fed to the fully connected layers, it is a common knowledge now that the activation vectors of fc-layers often serve as very good compact representations of the input signals (e.g., images).

2.3 Recurrent Neural Networks

Standard neural networks assume that input signals are independent of each other. However, often this is not the case. For instance, a word appearing in a sentence generally depends on sequence of the words preceding it. Recurrent neural networks (RNNs) are designed to model such sequences. An RNN can be thought to maintain a ‘memory’ of the sequence with the help of its internal states. In Fig. 2, we show a typical RNN that is unfolded-complete network is shown for the sequence. If the RNN has three layers, it can model e.g. sentences that are three words long. In the figure, x_t is the input at the t^{th} time stamp. For instance, x_t can be some quantitative representation of the t^{th} word in a sentence. The memory of the network is maintained by the state s_t that is computed as

$$s_t = f(Ux_t + Ws_{t-1}) \quad (5)$$

where $f(\cdot)$ is typically a non-linear activation function, e.g., ReLU. The output at a given time stamp is a function of a weighted version the network state at that time.

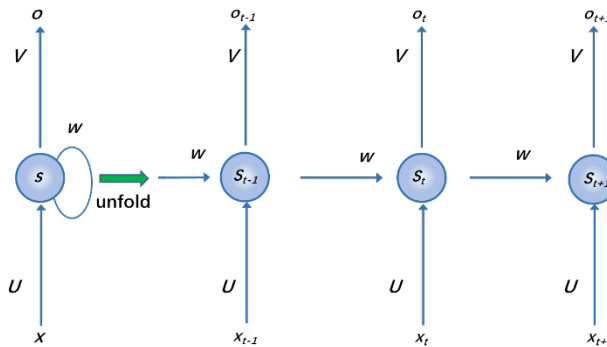


Figure 2: Schematic diagram of recurrent neural networks

3 Machine Learning Applications in Chemical Spectrum

In this section, we review the recent contributions in chemical spectrum analysis that exploit the machine learning technology. We mainly focus on the research papers published recent years, while briefly mentioning the more influential contributions from the earlier years. The literature pertaining to each task is then further sub-categorized based on the different spectrum.

3.1 X-Ray Spectra

For the anatomic region of X-Ray spectral, John, for the first time, analyzed X-ray absorption fine structure (XAFS) though ANN [20]. They mainly focus on using ANN method to build better models about complicated and multidimensional XAFS data of catalysis. It helps researchers understand the structure and function of catalysis and generate new knowledge about catalysis. Medford et al. [21] provided a hierarchy of data-information-knowledge derived from early works, utilizing machine-learning and uncertainty quantification.

Timoshenko et al. [22] presented supervised machine learning (SML) with input features of X-ray absorption near-edge structure (XANES) spectroscopy image dataset to polish the 3D metal catalysts structure. SML technique seeks to bridge the gap between the XANES fingerprints and catalyst geometry. Ab initio XANES simulations as the train data were used to train this SML model. Their experiment results reveal that SML could reconstruct the 3D geometry, including average size, shape, and morphology of well-defined nanoparticles (e.g., platinum). In the fact, there are essential difference between the computer-generated and experimental X-ray spectroscopy. However, the train dataset such as XANES and XAFS always use expensive and rare related instruments to gain. Extracting structural and electronic descriptors from XANES spectra is akin to solving a challenging inverse problem.

3.2 Raman Spectra

Raman spectrum can be regarded as a one-dimensional image. it can be solved by some kinds of neural network architecture. For instance, Lieber et al. [23] employed an automated method for fluorescence subtraction based on a modification to least-squares polynomial curve fitting. Because it is well-known that the inherent fluorescence of Raman spectrum generated by many biological molecules usually cover the true information among materials. Liu et al. [24] reported CNNs with multivariate treatment, including preprocessing, feature extraction and classification, to analyze Raman spectroscopy. This CNN method not only greatly simplifies the classification procedures for Raman spectroscopy, but also achieves high accuracy compared early related works using other machine learning methods such as support vector machine (SVM) with baseline corrected spectra. Chen et al. [25] described residual neural network (ResNet) was used to decode Raman spectra-encoded suspension arrays. This ResNet network presents classification stability and training convergence to different datasets, and reaches high classification accuracy 100%. Except the one-dimensional Raman images, Raman microscopic 2D image analysis in the context of bladder cancer cytopathology was also put forward by deep convolutional neural networks (DCNNs) applied in numerous pattern recognition tasks [26]. Going forward, it is significant to determine the detail information of chemical substance by automatic analysis of Raman spectra.

3.3 UV-Vis-IR Spectra

Baik et al. [27] demonstrated a system for identifying tablets using a portable visible-near-infrared (VIS-NIR) spectrometer and a CNN, which is one of the machine learning algorithms. According to spectroscopy techniques, the spectrum of each tablet has unique reflectance features. To classify tablets using their spectra, we have implemented three comparative experiments on the wavelength range and successfully classified 14 kinds of tablets. The results of the three experiments are 97.86% (using the VIS spectra), 96.90% (using the NIR spectra), and 98.81% (using the VIS-NIR spectra). This shows high accuracy without reference to the wavelength range. Bjerrum et al. [28] analyzed a tablet dataset using CNNs with automatic tuning of the model hyperparameters and regularization level in the form of dropout layers. The dataset consists of assay results from analysis of pharmaceutical tablets and NIR spectra recorded with two different instruments. Data preprocessing in the form of spectral data augmentation is

implemented and the performance compared with extended multiplicative scatter correction. All dataset treatments are compared to hypothetical optimal models as a baseline method. The model types performances are also compared on a specially crafted extrapolation challenge for both assay result values and instrument recordings.

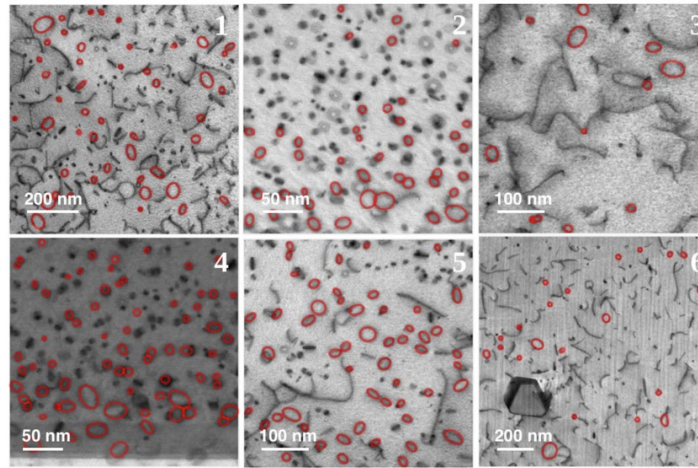


Figure 3: Fitted ellipse labeled by the automated machine learning program on the six selected test images. Image size: 1024×1024 pixels for all images except image 4, 2048×2048 pixels for image 4 [34]

3.4 Scanning Electron/Light Optical Microscope

The microstructure of material stores its genesis and determines all its physical and chemical properties. Since the microstructure could be a combination of different phases or constituents with complex substructures its automatic classification is very challenging. In assessing microstructures features, Velichko et al. [29] employed a method using data mining methods by extracting morphological features and a feature classification step on cast iron using SVMs. Jonathan present a max-pooling convolutional neural network approach for supervised steel defect classification [30]. Pauly et al. followed this same approach by applying on a contrasted and etched dataset of steel, acquired by scanning electron microscope (SEM) and light optical microscope (LOM) imaging using the steel images of Material Engineering Center Saarland (MECS) data set [31]. This dataset is available at the MECS website: <http://www.mec-s.de/>. However, it could only reach 48.89% accuracy in microstructural classification on the given dataset for four different classes due to high complexity of substructures and not discriminative enough features. More recently, Azimi et al. [32] proposed a deep learning method for microstructural classification in the examples of certain microstructural constituents of low carbon steel. This novel method employs pixel-wise segmentation via fully convolutional neural network accompanied by a max-voting scheme. Their system achieves 93.94% classification accuracy, drastically outperforming the state-of-the-art method of 48.89% accuracy. Beyond the strong performance of our method, this line of research offers a more robust and first of all objective way for the difficult task of steel quality appreciation. The authors used three possible CNN networks, namely CIFAR-Net proposed by Lecun et al. [19], AlexNet proposed by Alex et al. [33] and VGG19-Net [17], to categorize SEM and LOM images of steel. The accuracies reported by the authors for these classes are 51.27%, 56.56%, 60.02%, and 93.94%, corresponding to CIFARNet, VGG-19- DeCAF-RBF-kernel SVM, VGG-16 and MVFCNN respectively. As shown in Fig. 3, Li et al. [34] formulated an automated recognition tool based on a computer vision-based approach, including sequential a cascade object detector, convolutional neural network, and local image analysis methods. It indicates that this automated tool achieves high classification level in terms of recall and precision and achieves quantitative image/defect analysis metrics close to the human average. Katsumi et al. [35] proposed a new approach based on an image-processing or deep-learning-based method for super-resolution of 3D images with such asymmetric resolution, so as to restore the depth resolution to achieve symmetric resolution. The deep-learning-based method learns from high-resolution sub-images obtained via SEM and recovers low-

resolution sub-images parallel to the FIB milling direction. The 3D morphologies of polymeric nano-composites are used as test images, which are subjected to the deep-learning based method as well as conventional methods. they find that the former yields superior restoration, particularly as the asymmetric resolution is increased.

3.4 Photoluminescence

Wood et al. [36] proposed a machine learning algorithm to calculate time-resolved photoluminescence (TRPL) data and determine the probability distribution of decay rate an arbitrary emitter without any a priori assumptions. They employed Laplace transform and k-fold cross-validation to reduce the prediction error. In this work, Poisson statistics is used to solve the discrete Laplace transform, and the regularization method of k-fold cross validation is used to minimize the prediction error. For example, the attenuation rate distribution of Ag-In-Se nanocrystalline is best described by the machine learning model. This fitting method can be applied by researchers, especially to better understand the fluorescence principle of I-III-VI nanocrystalline. On the other hand, this machine learning model is also applied in energy-resolved, ultrafast emission from CsPbBr₃ perovskite nanocrystals. Besides, computer-generated TRPL data has been fitted and calculated by their method, indicating high practicability of this machine learning model. They provided their software (LumiML) open source [37]. The machine learning approach indicates the complex dynamics of neutral excitons, charged excitons and multiexcitons [38,39], as well as their binding energies, [40] highlighting the utility of this program for improving the comprehension of novel materials.

4 Future Directions

Currently, the number of chemical spectral works employed by machine learning has risen with the eruptible revolution of image analysis based on CNN. In the future, the academic area of chemical imaging analysis will follow an increasing large of research findings. Hence, we present core instructions and future directions to address the problems faced by machine learning in chemical spectra analysis. We provide our perspectives based on the related papers and other literature of similar interdisciplinary fields in machine learning/computer vision such as hyperspectral remote sensing data analysis, medical imaging analysis and so on. Therefore, we propose two following present challenges and corresponding solutions. (1) Owing to the lack of large-scale chemical imaging datasets, chemical spectral analysis will be limited to develop high-stability and enough-accuracy models. Fortunately, transfer learning can address the dilemma. Transfer learning owns the special ability of a system to recognize and apply knowledge and skills learned in other domains/tasks to novel domains/tasks [41]. Medical imagine analysis area has already employed this dramatic transfer learning [42–44]. It is obvious from this point that chemical spectra should make full use of transfer learning's applications in medical imaging field. (2) In the absence of large-scale chemical imaging datasets, employing the current machine learning model as high-level feature extractors carry out further training, which shows a brighter and promising direction in the future. By the way, the large-scale chemical spectral datasets should be developed with the need to not train computer-generated spectra, but experimental data.

5 Conclusion

In summary, we propose a survey of chemical spectral analysis via machine learning algorithm built upon the previous works. It generally presents the methodologies, applications and future directions. Fundamentally, due to the lack of understanding basic principles of machine learning among chemical spectral analyst, this review introduces the essential concepts of machine learning methods (e.g., supervised or unsupervised one) and deep learning frameworks (e.g., ANN, CNN, RNN). It can help chemical spectral analyst better understand the machine learning methods.

In the second section, it focuses on comprehensive machine learning uses in different chemical spectral. Unlike other existing related reviews in the field, this paper proposes computer vision/machine learning applications in the chemical imaging researchers' sense. This survey not only offers a deep understanding of the key concepts in machine learning for the chemical spectral field, but also puts emphasis on the intrinsic cause of the challenges in chemical spectral applications using machine learning.

Based on the recent literature, we point out that the deficiency of large chemical imaging datasets

occupies the main status in machine learning for chemical image analysis at the present time. Finally, we hold the view that scientists should adapt multiple measures to address this dilemma like other research areas via machine learning such as medical imaging analysis, hyperspectral remote sensing data analysis. We hope that this frontier interdisciplinary research area which combines chemical spectra and machine learning will raise scholars' interest among chemical analysts, computer vision and machine learning fields.

Funding Statement: This work is supported by National Natural Science Foundation of China (62072250).

Conflicts of Interest: We declare that we do not have any commercial or associative interest that represents a conflict of interest in connection with the work submitted.

References

- [1] E. Alpaydin, *Introduction to Machine Learning*, 3rd ed., vol. 1. Cambridge, MA, USA: MIT Press, pp. 5–19, 2009.
- [2] I. H. Witten, E. Frank, E and M. A. Hall, "Data mining: Practical machine learning tools and techniques," *ACM Sigmod Record Press*, vol. 31, no. 1, pp. 76–77, 2005.
- [3] N. Akhtar and A. Mian, "Threat of adversarial attacks on deep learning in computer vision: A survey," *IEEE Access*, vol. 6, pp. 14410–14430, 2018.
- [4] I. Sutskever and O. Vinyals, "Sequence to sequence learning with neural networks," arXiv preprint arXiv: 1409.3215, 2014.
- [5] S. Y. Kung, M. W. Mak and S. H. Lin, "Biometric authentication: A machine learning approach." Beijing, China: Pearson Education, pp. 45–72, 2004.
- [6] I. Kononenko, "Machine learning for medical diagnosis: history, state of the art and perspective," *Artificial Intelligence in medicine*, vol. 23, no. 1, pp. 89–109, 2001.
- [7] M. Imstaedter, K. Briggman, S. C. Turaga, V. Jain, H. S. Seung and W. Denk, "Connectomic reconstruction of the inner plexiform layer in the mouse retina," *Nature*, vol. 500, no. 7461, pp. 168–174, 2013.
- [8] T. Ciodaro, D. Deva and J. D. Seixas, "Online particle detection with neural networks based on topological calorimetry information," *Journal of Physics: Conference Series*, pp. 12–30, 2012.
- [9] H. Y. Xiong, B. Alipanahi, L. J. Lee, H. Bretschneider *et al.*, "The human splicing code reveals new insights into the genetic determinants of disease," *Science*, vol. 347, no. 6218, pp. 124, 2015.
- [10] A. Graves, A. Mohamed, and G. Hinton, "Speech recognition with deep recurrent neural networks," in *IEEE Int. Conf. on Acoustics, Speech and Signal Processing*, Florence, Italy, pp. 6645–6649, 2013.
- [11] L. Deng, "The MNIST database of handwritten digit images for machine learning research," *IEEE Signal Processing Magazine*, vol. 29, no. 6, pp. 141–142, 2012.
- [12] J. Ma, R. P. Sheridan, A. Liaw, G. E. Dahl and V. Svetnik, "Deep neural nets as a method for quantitative structure–activity relationships," *Journal of Chemical Information and Modeling*, vol. 55, no. 2, pp. 263–274, 2015.
- [13] R. J. Schalkoff, *Artificial Neural Networks*. McGraw-Hill Higher Education, 1997.
- [14] D. E. Rumelhart, G. E. Hinton and R. J. Williams, "Learning representations by back-propagating errors," *Cognitive Modeling*, vol. 5, no. 3, pp. 1, 1988.
- [15] D. P. Kingma and J. B. Adam, "A method for stochastic optimization," arXiv preprint arXiv:1412.6980 2014.
- [16] N. Qian, "On the momentum term in gradient descent learning algorithms," *Neural Networks*, vol. 12, no. 1, pp. 145–151, 1999.
- [17] A. Krizhevsky, I. Sutskever and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," *Advances in Neural Information Processing Systems*, vol. 25, no. 2, pp. 1097–1105, 2012.
- [18] O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, *et. al.*, "Imagenet large scale visual recognition challenge," *International Journal of Computer Vision*, vol. 115, no. 3, pp. 211–252, 2015.
- [19] C. Y. Le, L. Bottou, Y. Bengio and P. Haffner, "Gradient-based learning applied to document recognition," *Proceedings of the IEEE*, vol. 86, no. 11, pp. 2278–2324, 1998.
- [20] J. R. Kitchin, "Machine learning in catalysis," *Nature Catalysis*, vol. 1, no. 4, pp. 230, 2018.

- [21] A. J. Medford, M. R. Kunz, S. M. Ewing, T. Borders and R. Fushimi, "Extracting knowledge from data through catalysis informatics," *ACS Catalysis*, vol. 8, no. 8, pp. 7403–7429, 2018.
- [22] J. Timoshenko, D. Lu, Y. Lin, and A. I. Frenkel, "Supervised machine-learning-based determination of three-dimensional structure of metallic nanoparticles," *The Journal of Physical Chemistry Letters*, vol. 8, no. 20, pp. 5091–5098, 2017.
- [23] C. A. Lieber and A. Mahadevan-Jansen, "Automated method for subtraction of fluorescence from biological Raman spectra," *Applied spectroscopy*, vol. 57, no. 11, pp. 1363–1367, 2003.
- [24] J. Liu, M. Osadchy, L. Ashton, M. Foster, C. J. Solomon and S. J. Gibson, "Deep convolutional neural networks for Raman spectrum recognition: a unified solution," *Analyst*, vol. 142, no. 21, pp. 4067–4074, 2017.
- [25] X. Chen, L. Xie, Y. He, T. Guan and Y. Ji, "Fast and accurate decoding the Raman spectra encoded suspension array with deep learning," *Analyst*, vol. 144, no. 14, pp. 4312–4319, 2019.
- [26] S. D. Krauß, R. Roy, H. K. Yosef, T. Lechtonen, S. F. El-Mashtoly *et al.*, "Hierarchical deep convolutional neural networks combine spectral and spatial information for highly accurate Ramanspectroscopy-based cytopathology," *Journal of Biophotonics*, vol. 11, no. 10, pp. 145–159, 2018.
- [27] K. J. Baik, J. H. Lee, Y. Kim and B. J. Jang, "Pharmaceutical tablet classification using a portable spectrometer with combinations of visible and near-infrared spectra," in *Ninth Int. Conf. on Ubiquitous and Future Networks (ICUFN) IEEE*, Milan, Italy, pp. 1011–1014, 2017.
- [28] E. J. Bjerrum, M. Glahder and T. Skov, "Data augmentation of spectral data for convolutional neural network (CNN) based deep chemometrics," arXiv preprint arXiv:1710.01927 2017.
- [29] A. Velichko and F. Mücklich, "Quantitative 3D characterisation of graphite morphology in cast iron—correlation between processing, microstructure and properties," *International Journal of Materials Research*, vol. 100, no. 8, pp. 1031–1037, 2009.
- [30] J. Masci, U. Meier, C. Dan, J. Schmidhuber and G. Fricout, "In steel defect classification with max-pooling convolutional neural networks," in *Int. Joint Conf. on Neural Networks*, 2012.
- [31] J. Gola, D. Britz, T. Staudt, M. Winter, A. S. Schneider *et al.*, "Advanced microstructure classification by data mining methods," *Computational Materials Science*, vol. 148, no. 1, pp. 324–335, 2018.
- [32] S. M. Azimi, D. Britz, M. Engstler, M. Fritz and F. Mücklich, "Advanced steel microstructural classification by deep learning methods," *Scientific Reports*, vol. 8, no. 1, pp. 21–28, 2018.
- [33] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," arXiv preprint arXiv:1409.1556 2014.
- [34] W. Li, W. K. G. Field and D. Morgan, "Automated defect analysis in electron microscopic images," *NPJ Computational Materials*, vol. 4, no. 1, pp. 36, 2014.
- [35] K. Hagita, T. Higuchi and H. Jinnai, "Super-resolution for asymmetric resolution of FIB-SEM 3D imaging using AI with deep learning," *Scientific Reports*, vol. 8, no. 1, pp. 58–77, 2018.
- [36] N. Đorđević, J. S. Beckwith, M. Yarema, O. Yarema, A. Rosspeintner *et al.*, "Machine learning for analysis of time-resolved luminescence data," *ACS Photonics*, vol. 5, no. 12, pp. 4888–4895, 2018.
- [37] <http://www.lumiml.ethz.ch/>.
- [38] D. Canneson, E. V. Shornikova, D. R. Yakovlev, T. Rogge, A. A. Mitoglu *et al.*, "Negatively charged and dark excitons in CsPbBr₃ perovskite nanocrystals revealed by high magnetic fields," *Nano Letters*, vol. 17, no. 10, pp. 6177–6183, 2017.
- [39] M. Fu, P. Tamarat, H. Huang, J. Even, A. L. Rogach *et al.*, "Neutral and charged exciton fine structure in single lead halide perovskite nanocrystals revealed by magneto-optical spectroscopy," *Nano letters*, vol. 17, no. 5, pp. 2895–2901, 2017.
- [40] J. A. Castañeda, G. Nagamine, E. Yassitepe, L. G. Bonato, O. Voznyy *et al.*, "Efficient biexciton interaction in perovskite quantum dots under weak and strong confinement," *ACS Nano*, vol. 10, no. 9, pp. 8603–8609, 2016.
- [41] Q. Yang, "An Introduction to Transfer Learning," in *Advanced Data Mining and Applications*, vol. 5139. Berlin, Heidelberg: Springer, 2008.
- [42] S. Azizi, P. Mousavi, P. Yan, A. Tahmasebi, J. T. Kwak *et al.*, "Transfer learning from RF to B-mode temporal enhanced ultrasound features for prostate cancer detection," *International Journal of Computer Assisted*

Radiology and Surgery, vol. 12, no. 7, pp. 1111–1121, 2017.

- [43] T. Kooi, B. Ginneken, N. Karssemeijer and A. den Heeten, “Discriminating solitary cysts from soft tissue lesions in mammography using a pretrained deep convolutional neural network,” *Medical Physics*, vol. 44, no. 3, pp. 1017–1027, 2017.
- [44] R. K. Samala, H.P. Chan, L. M. Hadjiiski, M. A. Helvie, K. H. Cha *et al.*, “Multi-task transfer learning deep convolutional neural network: Application to computer-aided diagnosis of breast cancer on mammograms,” *Physics in Medicine & Biology*, vol. 62, no. 23, pp. 88–94, 2017.