Applied AI in Instrumentation and Measurement: The Deep Learning Revolution

Mounib Khanafer and Shervin Shirmohammadi

In the last few years, hardly a day goes by that we do not hear about the latest advancements and improvements that Artificial Intelligence (AI) has brought to a wide spectrum of domains: from technology and medicine to science and sociology, and many others. AI is one of the core enabling components of the fourth industrial revolution that we are currently witnessing, and the applications of AI are truly transforming our world and impacting all facets of society, economy, living, working, and technology. The field of Instrumentation and Measurement (I&M) is no exception, and has already been impacted by Applied AI. In this article, we give an overview of Applied AI and its usage in I&M. We then take a deeper look at the I&M applications of one specific AI method: Deep Learning (DL), which has recently revolutionized the field of AI. Our survey of DL papers published in the *IEEE Transactions on Instrumentation and Measurement* (IEEE TIM) and *IEEE Instrumentation & Measurement Magazine* showed that, since 2017, there is a very strong interest in applying DL methods to I&M, in terms of measurement, calibration, and other I&M challenges. In particular, of the 32 surveyed papers, 75% were published in 2017 or later, and a remarkable 50% were published in 2019 alone. Considering that 2019 was not yet finished when we were writing this article, the recent exponential interest in and impact of DL in I&M is a very evident trend. We also found that although DL is used in a variety of I&M topics, a considerable portion of DL in I&M focuses on Vision Based Measurement (VBM) systems (around 28%) and fault/defect diagnosis/detection/prediction (around 25%). Finally, we found that Convolutional Neural Networks are the most widely used DL technique in I&M, especially in VBM. But to explain all of the above findings, we first need to understand AI itself and what we mean by it in its applied context. So let us begin our discussion with Applied AI.

Applied Artificial Intelligence

Although the long-term goal of research in AI is to enable machines to have the same level of intelligence as animals or humans, it is important to note that the Applied AI of today is not really comparable to biological intelligence. In fact, the word *Intelligence* in the Applied AI of today is misleading for the common person, as it gives the person the wrong impression that he or she is dealing with an intelligent being manifested as AI. This becomes even more misleading if we consider that we do not even have a universally agreed-upon definition of *intelligence*! In other words, experts do not yet completely understand or agree what intelligence is. So, then, how can we call something *artificial intelligence* if we do not even know what *intelligence* is? This has been an ongoing debate among AI experts for a long time. For brevity, in this article we do not enter this debate, and we refer the interested readers to other sources such as [1].

So, what does it mean when we talk about Applied AI in today's context, if it does not mean biological intelligence? Today, Applied AI practically means the application of AI *methods* as tools to advance or improve a system in a given domain; for example using AI methods to improve the weather forecasting system in meteorology, increase the efficiency of warehouse logistics in the storage and shipping industry, obtain earlier diagnosis of diseases in medicine, or in the I&M domain, reduce complexity of a measurement method or increase accuracy of a measurement instrument. Researchers expect that, in the future, these AI methods will match or even surpass human intelligence. But today we are neither close to that goal, nor do we need to be close in the applied arena, because AI methods as they are today are already having a beneficial impact on existing systems.

So, what are these methods? Fig. 1 shows some of AI's most commonly applied methods, which are normally inspired by how the biological brain works, referred to as *computational intelligence* [2]; for example, artificial neural networks try to emulate the neural networks in a biological brain, or *fuzzy logic* tries to operate similar to how humans make decisions without complicated mathematical modeling and using only imprecise or vague information. What has made Applied AI such a disruptive technology is the fact that these methods allow us to do tasks such as classification, clustering, prediction, decision making, and optimization *without* the need to first build an analytical model of the problem or the system at hand. This is very important, so let us take a closer look at this.

1094-6969/20/$25.00©2020IEEE

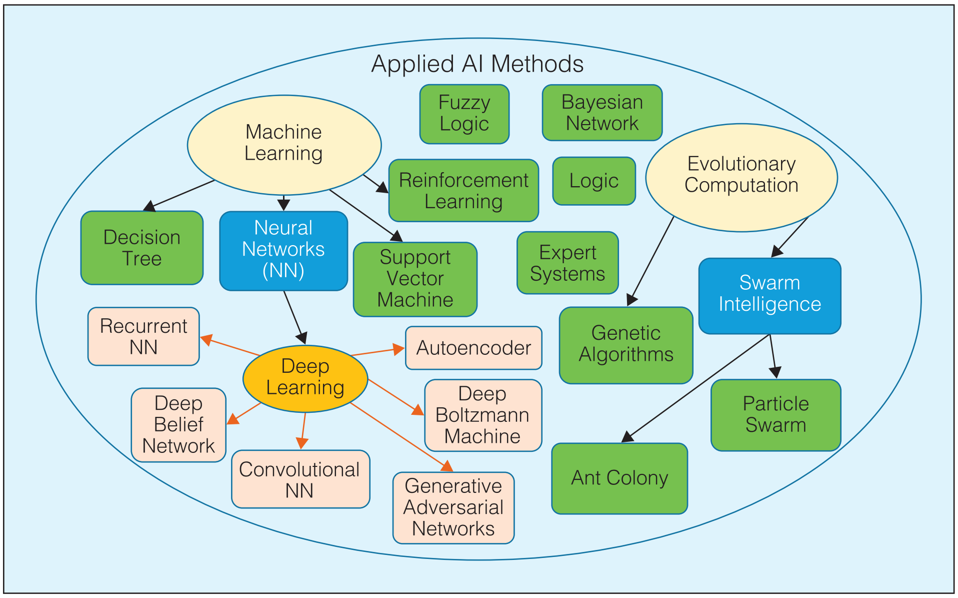

Fig. 1. Some of the most-commonly used Applied AI methods. Deep Learning, the subject of this article, is indicated in orange. Note that the methods shown in the figure are not exhaustive.

Most of these methods take as input a set of previously gathered data and try to come up with a model that best matches this input data to the desired output. This model can then be used to obtain output values for new incoming data. This means that the model that an AI method creates is purely based on matching input to output (black box), unlike the analytical model that a domain expert creates (white box). For example, let us say we want to create a system that detects skin cancer from the image of a patient's skin lesions. To create such a system analytically, a biomedical engineer needs to develop an analytical model based on the skin lesions' shape, pattern, brightness, colour, concentration, area, and many other parameters, some of which are complicated to model, impossible to model, or simply unknown. Development of such an analytical model could be very complex, especially if the model needs to be generalizable; i.e., working for any random skin type or color. There are too many parameters here to come up with a general model that would work with acceptable accuracy. But, if we use an AI method, for example a machine learning algorithm, we only need to train it with a large-enough dataset of previously taken images of skin lesions with and without cancer. The method then creates its own model of how to match features in those images to whether or not cancer is present. This not only significantly reduces complexity, but also in some cases gives an even better result than an analytical model. Of course, the accuracy of this AI-based model depends on the specific AI method and algorithm selected, as well as the quality of the provided dataset. With this explanation of Applied AI in mind, let us now take a look at how we can use it in I&M.

Applied AI in I&M

Looking at the scope webpage of *IEEE Transactions on Instrumentation & Measurement* (IEEE TIM), (http://tim.ieee-ims.org), we can say that I&M as a research field is interested in contributions to "methods or instruments for measurement, detection, tracking, monitoring, characterization, identification, estimation, or diagnosis of a physical phenomenon; uniqueness of an application furthering the I&M fields; or measurement theory including uncertainty, calibration, etc." The popularity of AI in I&M is therefore not surprising, considering that AI methods are quite suitable for detection, tracking, monitoring, characterization, identification, estimation, diagnosis, or prediction. In fact, the huge potential of using AI in I&M was already noticed by the 1990s, for example in [3], which describes very well neural networks, fuzzy logic, and expert systems, plus their applications in I&M, such as sensor design, calibration, uncertainty prediction, measurement data interpolation, software instruments, indirect measurement, fault detection, and system identification. AI is especially practical when the precise and accurate mathematical modeling of the measurement system or instrument at hand is highly complex, highly nonlinear, highly dynamic, or impossible due to lack of knowledge of the system except for a limited number of parameters. In such cases, or when the accuracy of the final measurement is more important than understanding exactly how the system works, AI can offer an attractive and practical solution. One example that well illustrates the above notion is Vision-Based Measurement (VBM) [4], which requires AI as a core component. An example of a VBM system is [5], which indirectly measures the amount of calories and nutrition in food from the food's picture. Such a system would simply not be implementable without AI, due to the highly complex nature of the problem and lack of complete knowledge of all parameters. Another area where AI can be applied is calibration, for example in [6], where a picture of the user's hand is used to easily calibrate a complex force measuring instrument. Measurement prediction is another useful feature of AI, which can be applied when actual measurement is either costly or impractical. For example, in a massively multisource networked system, it is not possible to explicitly measure the end-to-end delay between all pairs of nodes on the network, due to the $O(N^2)$ complexity of the problem. In such a case, AI has been shown to predict measurements more accurately and orders of magnitude faster than non-AI techniques [7].

In Fig. 1 we can also see Deep Learning (DL), the main subject of this article. Although DL has existed since the 1980s [8], it wasn't until 2011 and 2012 that it caused significant publicity among researchers and practitioners, when DL-based visual object recognition systems such as IDSIA and AlexNet beat out their competition by huge margins in terms of improved error rate [9], thanks partly to GPUs providing the significant speed up needed by DL. Another reason for their popularity is that DL algorithms automate the process of extracting discriminative features in a dataset, which usually requires domain knowledge and significant human efforts [10]. This is why DL has revolutionized Applied AI in recent years. To understand how DL can achieve such great results, and how it is being used in I&M, let us now take a look at the basics of DL.

Deep Learning Basics

A basic understanding of neural networks, as described in [3], can be helpful to the readers here. A DL algorithm is based on a layered architecture of data representation in which the high-level features are extracted from the last layers of the neural network while the low-level features are extracted from the lower layers [11]. The true inspiration behind such an architecture is to mimic how the biological brain works—as brains extract data representation from an input, e.g., scene information from the eyes, the generated output is a classified object [11].

One main capability of DL is to extract complex features from a huge amount of data and discover hidden patterns and trends in them. This is achieved by the utilization of a neural network that is constituted by a set of interconnected neurons (computational or processing units). A single neuron receives data from inputs or other neurons, multiplies them with weights and then feeds them through an activation function to produce an output as depicted in Fig. 2.

In a neural network, neurons are organized into an input layer, one or more hidden layers, and an output layer as shown in Fig. 3. The input layer is composed of input neurons that feed in the input data. A hidden layer receives its input data either from the input layer neurons or the neurons of a preceding hidden layer. Hidden layers perform intensive processing of the data that were originally supplied by the input layer. Apparently, the more hidden layers the neural network has, the *deeper* and more intense processing the data go through. In case the neural network architecture is designed with two or more hidden layers, we end up with a DL architecture. Definitely, the use of more hidden layers entails the need for higher computational and processing capability, an area that GPUs can assist tremendously. Finally, the output layer produces the final output of the neural network.

How a DL algorithm learns about features and manages to extract them depends on the availability of the training dataset. Like any machine learning technique, the DL approach of learning can be done as either supervised learning or unsupervised learning. In supervised learning, the dataset is comprised of labeled data. That is, we deal with input data for which the respective output is known and defined. In unsupervised learning, however, we deal with unstructured data for which the output is unknown.

Training the DL network with the training dataset works as follows. As data flows from the input layer to the first hidden layer, the neurons in this layer use their activation functions to

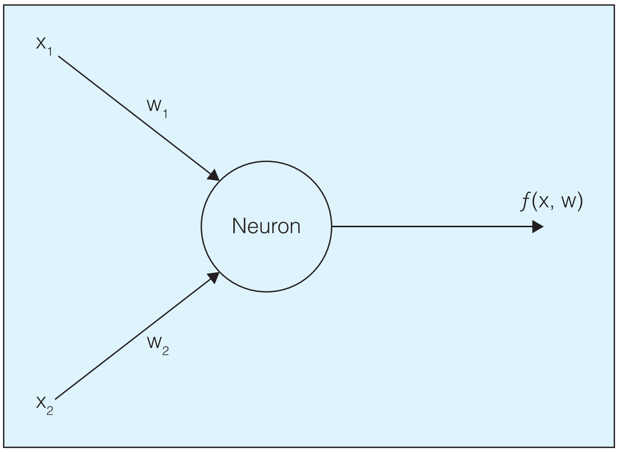

Fig. 2. A depiction of the operation of a neuron: it receives inputs (like x_1 and x_2), multiplies them with weights (like w_1 and w_2), sums up the multiplication results, and finally passes the sum to an activation function that produces the output $f(x,w)$.

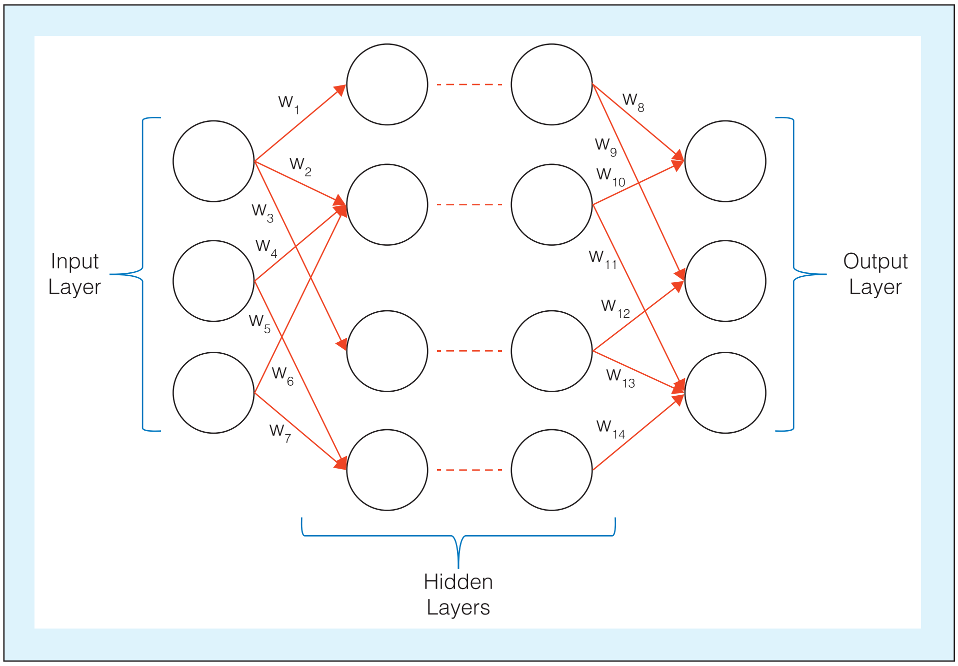

Fig. 3. Architecture of a neural network. The W_1, W_2, ... , W_{14} parameters are the weights of the connections.

produce outputs, as explained earlier, that are fed into the second hidden layer. The latter repeats what the previous hidden layer performed and feeds its outputs as inputs into the next hidden layer. This process is repeated until the output layer produces the final output. If the final output differs from the foreseen output, the weights on the connections between the layers are adjusted and the whole process is repeated starting from the input layer. In each iteration of training, the amount of deviation between the generated output and the expected output is measured; this is referred to as the loss function. The ultimate goal is to bring the loss function as close to zero as possible, making the DL algorithm more accurate in extracting the correct features or patterns from a given set of input data. But this also means that DL inherently has some level of uncertainty, which we discuss next.

Uncertainty in Deep Learning

DL applications vary in the level of uncertainty they can tolerate. While we can tolerate most Siri mistakes that appear regularly, critical applications like autonomous driving or content filtering for copyright legislation require models with much less uncertainty in their predictions. Therefore, let us shed some light on the uncertainty in DL techniques and how it is addressed by researchers.

Like any other measurement system, two types of effects can contribute to uncertainty in a DL system, namely: systematic and random [12]. Systematic effects occur when the DL system is trained with a dataset that is not sufficiently representative of the entire input domain, or if the DL model underfits the training data. Even if a DL model does not underfit the training data, it can reasonably predict measurement results based on data patterns that it has been trained with before; uncertainty may rise when it is fed with data patterns that it is seeing for the first time. This can be improved by training the DL with more generalizable data, and with using better hypothesis functions. Random effects, on the other hand, arise if during prediction (not during training) the DL model slightly changes, such as with Monte Carlo Dropout, which causes the DL model to produce different outputs for the same input depending on the nodes that are dropped. These random effects will be larger if the DL model suffers from overfitting during training. Whether we are dealing with systematic or random effects, it is essential to quantify the level of uncertainty associated with the DL techniques. In DL, we can typically use the following methods to quantify the uncertainty [13]: Monte-Carlo Dropout, Deep Ensembles (e.g., Distributional Parameter estimation or Ensemble Averaging), Dropout Ensembles, Quantile Regression, and Gaussian Process Inference. Details are beyond the scope of this article, and in fact we are writing an article just about uncertainty in DL-based measurements which we hope will be published in this same magazine in the near future. With the above general description about DL and its uncertainty in mind, let us now take a look at specific DL techniques used in practice.

Deep Learning Techniques

The DL techniques that we observed in our literature search include recurrent neural networks, convolutional neural networks, deep Boltzmann machines, deep belief networks, generative adversarial networks, and autoencoders.

Recurrent Neural Networks

Recurrent neural networks (RNNs) are developed to process sequential information. In these networks, the same task is applied for every element in the sequence of information, with the output being dependent on the previous computations [14]; thus, the name *recurrent*. The typical architecture of an RNN is illustrated in Fig. 4a. In this figure, N designates part of a neural network, X_t refers to inputs over time, and Y_t refers to the corresponding outputs over time.

It is quite convenient to unfold the architecture in Fig. 4a to look like the one in Fig. 4b. In this figure, we can clearly see that a type of memory is employed in RNNs where the information that has been calculated till time *t* is preserved. One popular model for an RNN is the long short-term memory (LSTM) model that was proposed by Gers and Schmidhuber in [15]. Typical applications that benefit from RNNs are speech recognition, language modeling, translation, and image captioning [14].

Convolutional Neural Networks

Convolutional neural networks (CNNs) employ convolutional layers, subsampling (pooling) layers, and a final stage of a fully connected layer, as illustrated in Fig. 5. These networks are named *convolutional* due to their use of the convolution mathematical operation in their feature extraction process.

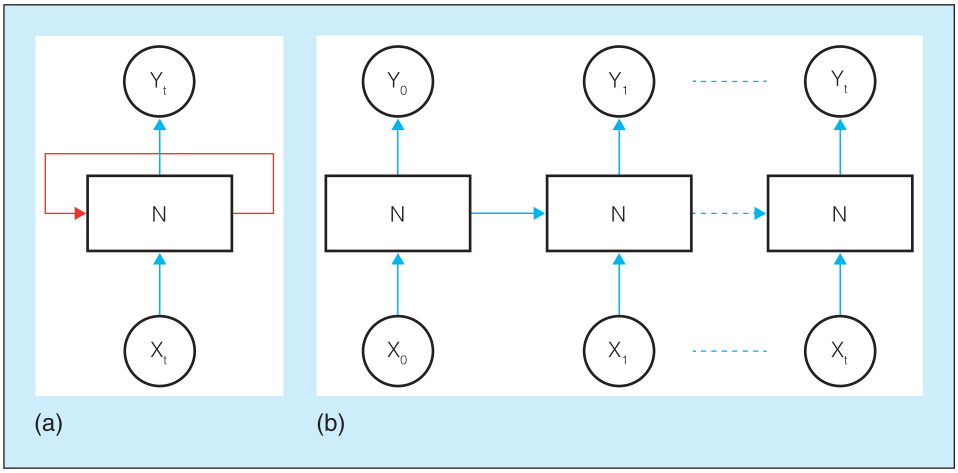

Fig. 4. a) The typical architecture of an RNN. b) Unfolded RNN architecture.

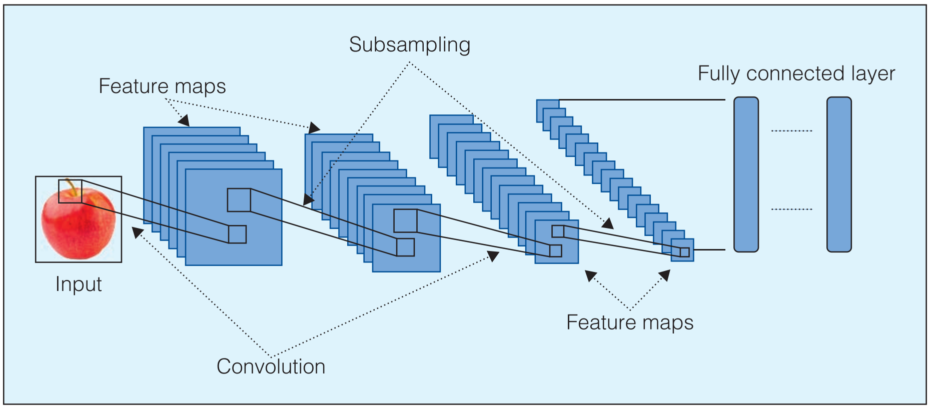

Fig. 5. Typical architecture of a CNN.

Both convolution and subsampling layers are used for feature extraction. Specifically, as shown in Fig. 5, convolution layers produce feature maps while subsampling layers reduce the sizes of these maps without losing the key information in them. The final output of these layers is fed into the fully connected layer which handles the task of classification. CNNs are typically used in computer vision and speech recognition applications [11].

Deep Boltzmann Machines and Deep Belief Networks

In a Boltzmann machine, nodes are fully connected to each other using undirected edges, as shown in Fig. 6a. The nodes are divided into visible and hidden nodes, with no output layer of nodes. In restricted Boltzmann machines (RBMs), however, the visible nodes are only connected to hidden nodes, and vice versa. Deep Boltzmann machines (DBMs) are formed by stacking RBMs on top of each other. The hidden nodes form full connectivity across subsequent layers. The visible nodes are the ones that receive the training set of data. Compared to other DL networks, DBMs are unique in that the input visible nodes are connected to each other. The ultimate goal of a DBM is to detect the distribution of the given input training dataset. Typical applications of DBMs are speech and object recognition [11].

Deep belief networks (DBNs) are based on stacking RBM layers on top of each other. However, while top layers have undirected edges, lower layers have directed ones, as shown in Fig. 6b.

Generative Adversarial Networks

Generative adversarial networks (GANs) suggest an approach for maximum likelihood estimation and employ two neural networks that compete against each other in a zero-sum game [16]. Basically, the GAN architecture utilizes a *generator* model and a *discriminator* model. The generator is the system that is trained to generate images, while the discriminator is the system that classifies these images as accurate or not. This process of image generation and classification is repeated until the generator is able to produce accurate results. Examples of applications of GANs include game development and artificial video generation [16].

Autoencoder

Autoencoding is a DL algorithm that efficiently compresses information and learns how to reproduce accurate approximation of the original information from the compressed data. Since they are used for data compression, autoencoders are efficient tools for dimensionality reduction. Typical where autoencoders are applied include data denoising and dimensionality reduction for data visualization [14].

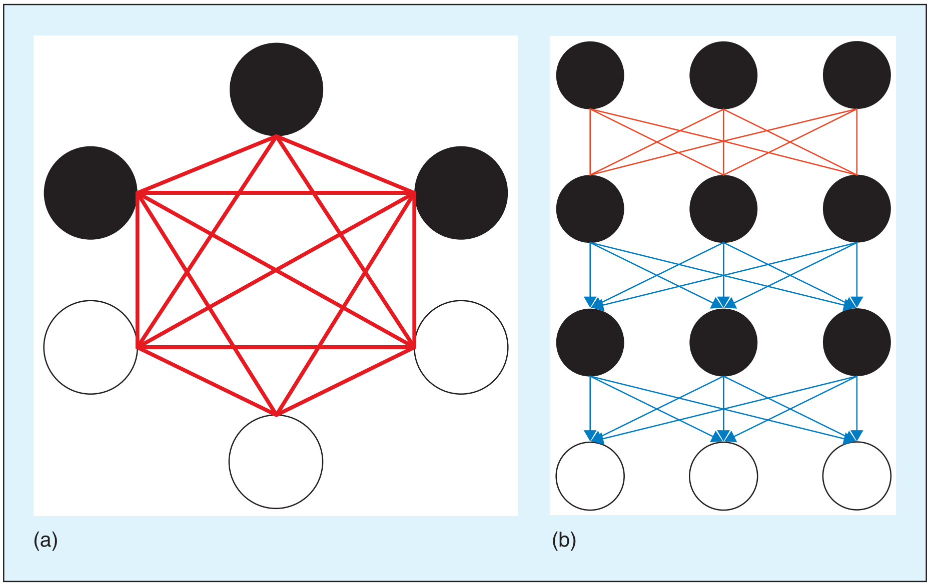

Fig. 6. a) Deep Boltzmann machine (DBM). The black nodes are hidden neurons while the white nodes are visible neurons. b) Architecture of the deep belief network (DBN). The black nodes are hidden neurons and organized into several layers while the white nodes are visible neurons. Lower layers use directed edges as opposed to fully undirected edges in DBMs.

Deep Learning in Instrumentation and Measurement

Here, we describe examples of how DL is being used in I&M literature, with a note that our search was inexhaustive and mostly limited to DL papers appearing in

IEEE TIM. We categorize the examples based on the DL technique used.

CNN

CNN has been used extensively in VBM systems. For example, a three-stage automatic defect inspection system for high-speed railways has been proposed in PVANET++ [17]. As split pins have a key role in fixing joint components on catenary support devices (CSDs), PVANET++ localizes and inspects split pins by using CNN to detect defects. Another system uses a two-stage deep CNN to detect insulator surface defects in railway catenary [18] by firstly localizing the catenary components to obtain images of the insulators, and then using a deep material classifier and a deep denoising autoencoder to detect defects. Since the defect detection of the fasteners on a CSD is essential for both safety and cost reductions in the operation of high-speed railways, a VBM method based on a deep CNN [19] is used to tackle the design of such detection systems.

A VBM system for the purpose of measuring pain intensity through the analysis of facial expressions is presented in [20] which uses the aforementioned AlexNet architecture to extract critical features from patients' images and to draw conclusions on the level of pain they are experiencing. The work in [21] proposes a vision-based evaluation (VBE) framework for VBM systems that studies the capability of DL strategies to deal with uncertainty contributions, a very important issue in I&M. This framework is also centered around an AlexNet-based CNN and handles uncertainty contributions during calibration processes. Another VBM system is presented in [22] which uses both CNN and DBM to design a food recognition system that analyzes pictures of meals, taken by a mobile device, to identify different food items and estimate their calories and nutrition. Yet another VBM system [23] tackles thermal image processing. A dataset of high-resolution thermal facial images is built and used to train the deep alignment network (DAN) algorithm (which is based on CNN) for face analysis. CNN is also used in robotic VBM systems, for example [24] uses a Faster Region CNN to aid robots in the recognition of hand gesture, while [25] introduces a projection algorithm to generate RGB, depth, and intensity (RGB-DI) images to measure the outdoor environments with a variable resolution. A full CNN is then used to segment those RGB-DI images and use them to realize the 3-D scene understanding for mobile robots.

The application of DL principles in light detection and ranging (LiDAR)-based perception tasks is also studied, for example, in [26]. This study, due to the limited availability of LiDAR point cloud datasets, uses simulators to automatically generate 3-D annotated LiDAR point clouds, and then uses that data to train a deep model that incorporates both CNN and RNN principles. Since network binarization is associated with reductions in computational and memory costs in 2-D computer vision tasks, [27] proposes a technique to train binary volumetric CNNs for 3-D object recognition.

CNN has also been used for extracting features from vibration signals to aid in the diagnosis of faults of a magnet synchronous motor [28], fault diagnosis of spindle bearings [29], and safety of the pipelines used for liquid petroleum transportation [30]. As magnetic flux leakage (MFL) is a common testing method for those purposes, the defects are identified from MFL images based on CNNs.

RNN

RNNs have been used in the detection of fundamental phasors and the identification of control and protection signals in power systems [31], as well as estimating lithium-ion battery remaining useful life to achieve an intelligent battery management system [32], and supporting predictive models for voltage correction in a dc voltage reference source [33]. In nonlinear dynamical systems, RNNs are used to tackle the identification problem: finding a time-dependent model for the behavior of the process generating the data [34], as well as in the modelling of dynamic systems in the absence of measurable state variables [35]. Another use of RNNs is in modeling of processes that occur in various industries like steam-raising plants, gas turbines, and automotive suspensions [36], since the output of such processes has unsymmetrical behavior. Finally, RNNs have been used for fault detection: in wireless sensor networks for sensor node fault detection [37], health condition monitoring of a machine [38], or early fault detection in industrial applications [39], where a deep neural network is used for feature extraction and the LSTM network for distribution estimator.

DBN

A discriminative DBN and ant colony optimization model has been built in [40] for the purpose of health condition monitoring of machines. DBN is also used to distinguish acceptable and unacceptable segments in an electrocardiogram (ECG) signal [41], to reduce false alarms during atrial fibrillation detection. Finally, DBN has been used in a soft sensor for estimating the deflection of a polymeric mechanical actuator [42]. The latter is based on ionic polymer-metal composites, which are used in important fields like robotics and surgery.

Autoencoders

Autoencoders are used in rotating machinery measurements, for example in fault identification for rotating machines [43] using a stacked sparse autoencoders-based deep neural network coupled with the concept of compressed sensing, or in studying the conditions of rotating machinery [44], where sensors generate fault signals, and features are extracted from these signals and fused by sparse autoencoder (SAE) neural networks. The fused features are then classified by a DBN. Another work that uses autoencoders for fault diagnosis is based on the analysis of vibration signals in wind turbine gearboxes [45], where stacked multilevel-denoising autoencoders are used to assist in learning discriminative representations of fault features.

Localization is another application of autoencoders. For example, the human step length in Pedestrian Dead Reckoning systems can be estimated with stacked autoencoders [46].

Finally, deep autoencoders are used for feature extraction to mitigate the corruption of Photoplethysmographic (PPG) measurements due to motion artifacts in personal healthcare systems [47].

Generative Adversarial Network (GAN)

GAN is one of the newest members of the DL family, and so it is not surprising not to see as many applications with GAN as with the other DL techniques. Among the papers we studied, GAN was used to measure the reliability of transmission gears [48].

Observations and Future Trends

The following observations can be made from the studied literature:

- The spectrum of applications benefiting from DL is quite broad. Researchers from various fields (industrial and systems engineering, operations, food control, health care, machinery, transportation, image processing, circuit design, sensor networks, etc.) are benefiting from DL architectures. This ought to drive DL research in I&M forward and pave the way for promising solutions to known and/or unresolved issues.
- Several papers stress that their proposed solutions are the *first* to apply DL to their problem of interest. This popularity of using DL is stimulated by the fact that an abundance of datasets is becoming available for training DL algorithms. In earlier years, many applications suffered from the scarcity of the available training data.
- There is a strong focus on using CNN. The reason is that CNNs are proven to be efficient in computer vision applications [12], so they are the natural first choice in VBM systems. The fact that rich datasets of images and videos have been recently made available for training purposes boosts the usage of AI and of CNN in particular.
- A considerable percentage of the publications focus on VBM (around 28%) and on fault/defect diagnosis/detection/prediction (about 25%). Many industry sectors can highly benefit from such research.

Despite the above encouraging findings, two clear gaps were observed in the limited literature that we searched:

- The usage of DL in biometric and security systems: DL can be an excellent choice to design face recognition, identity check, or other types of user identification components for security systems.
- As Industry 4.0 unfolds, it appears as a major area to leverage DL. The industrial Internet of Things (IoT) paradigm in particular appears to be a strong candidate for the utilization of DL architectures. Industrial IoT is heavily involved in monitoring activities, calibration and control of sensor nodes, fault detection, etc., which can benefit from the capabilities of DL. Also, given the fact that a huge amount of data is generated and processed within an IoT platform, autoencoders can be an efficient tool to compress data, which reduces the burden on the communication system.

References

[1] K. Warwick and S. J. Nasuto, "Historical and current machine intelligence," *IEEE Instrum. Meas. Mag.*, vol. 9, no. 6, pp. 20-26, Dec. 2006.

[2] D. B. Fogel, "Machine intelligence: teaching computers to learn," *IEEE Instrum. Meas. Mag.*, vol. 9, no. 3, pp. 12-16, Jun. 2006.

[3] C. Alippi *et al.*, "Artificial intelligence for instruments and measurement applications," *IEEE Instrum. Meas. Mag.*, vol. 1, no. 2, pp. 9-17, Jun. 1998.

[4] S. Shirmohammadi and A. Ferrero, "Camera as the instrument: the rising trend of vision based measurement," *IEEE Instrum. Meas. Mag.*, vol. 17, no. 3, pp. 41-47, Jun. 2014.

[5] P. Pouladzadeh *et al.*, "Measuring calorie and nutrition from food image," *IEEE Trans. Instrum. Meas.*, vol. 63, no. 8, pp. 1947-1956, Aug. 2014.

[6] J. Zhou *et al.*, "A new hand measurement method to simplify calibration in CyberGlove-based virtual rehabilitation," *IEEE Trans. Instrum. Meas.*, vol. 59, no. 10, pp. 2496-2504, Oct. 2010.

[7] S. A. Mohammed *et al.*, "A multimodal deep learning based distributed network latency measurement system," *IEEE Trans. Instrum. Meas.*, Jan. 2020.

[8] R. Dechter, "Learning while searching in constraint-satisfaction problems," University of California, in *Proc. AAAI National Conf. Artificial Intelligence*, pp. 178-183, Aug. 1986.

[9] J. Schmidhuber, "History of computer vision contests won by deep CNNs on GPU," March 2017. [Online]. Available: http://people.idsia.ch/~juergen/computer-vision-contests-won-by-gpu-cnns.html.

[10] M. M. Najafabadi *et al.*, "Deep learning applications and challenges in Big Data analytics," *J. Big Data*, vol. 2, no. 1, pp. 1-21, Feb. 2015.

[11] S. Pouyanfar *et al.*, "A survey on deep learning: algorithms, techniques, and applications," *ACM Computing Surveys*, vol. 51, no. 5, Article 92, Sep. 2018.

[12] A. Ferrero and S. Salicone, "Measurement uncertainty," *IEEE Instrum. Meas. Mag.*, vol. 9, no. 3, pp. 44-51, Jun. 2006.

[13] S. Bachstein, "Uncertainty quantification in deep learning," Master's Thesis, Ulm University, Germany, Jan. 2019.

[14] M. Coşkun *et al.*, "An overview of popular deep learning methods," *European J. Technic.*, vol. 7, no. 2, pp. 165-176, Dec. 2017.

[15] F. A. Gers and J. Schmidhuber, "Recurrent nets that time and count," in *Proc. IEEE-INNS-ENNS Int. Joint Conf. Neural Networks*, vol. 3, Jul. 2000.

[16] M. Z. Alom *et al.*, "A state-of-the-art survey on deep learning theory and architectures," *Electronics*, vol. 8, no. 292, pp. 1-67, Mar. 2019.

[17] J. Zhong *et al.*, "A CNN-based defect inspection method for catenary split pins in high-speed railway," *IEEE Trans. Instrum. Meas.*, vol. 68, no. 8, pp. 2849-2860, Aug. 2019.

[18] G. Kang *et al.*, "Deep architecture for high-speed railway insulator surface defect detection: denoising autoencoder with learning," *IEEE Trans. Instrum. Meas.*, vol. 68, no. 8, pp. 2679-2690, Aug. 2019.

[19] J. Chen *et al.*, "Automatic defect detection of fasteners on the catenary support device using deep convolutional neural network," *IEEE Trans. Instrum. Meas.*, vol. 67, no. 2, pp. 257-269, Feb. 2018.

[20] P. Casti *et al.*, "Calibration of vision-based measurement of pain intensity with multiple expert observers," *IEEE Trans. Instrum. Meas.*, vol. 68, no. 7, pp. 2442-2450, Jul. 2019.

[21] G. Callari *et al.*, "The influence of uncertainty contributions on deep learning architectures in vision-based evaluation systems," *IEEE Trans. Instrum. Meas.*, vol. 68, no. 7, pp. 2425-2432, Jul. 2019.

[22] P. Pouladzadeh and S. Shirmohammadi, "Mobile multi-food recognition using deep learning," *ACM Trans. Multimedia Computing, Commun. Applications*, vol. 13, no. 3s, article 36, Aug. 2017.

[23] M. Kopaczka *et al*., "A thermal infrared face database with facial landmarks and emotion labels," *IEEE Trans. Instrum. Meas*., vol. 68, no. 5, pp. 1389-1401, May 2019.

[24] C. Nuzzi *et al.,* "Deep learning-based hand gesture recognition for collaborative robots," *IEEE Instrum. Meas. Mag*., vol. 22, no. 2, pp. 44-51, Apr. 2019.

[25] Z. Qiu *et al*., "RGB-DI images and full convolution neural network-based outdoor scene understanding for mobile robots," *IEEE Trans. Instrum. Meas*., vol. 68, no. 1, pp. 27-37, Jan. 2019.

[26] F. Wang *et al*., "Automatic generation of synthetic LiDAR point clouds for 3-D data analysis," *IEEE Trans. Instrum. Meas*., vol. 68, no. 7, pp. 2671-2673, Jul. 2019.

[27] C. Ma *et al*., "Binary volumetric convolutional neural networks for 3-D object recognition," *IEEE Trans. Instrum. Meas*., vol. 68, no. 1, Jan. 2019.

[28] I-H. Kao *et al*., "Analysis of permanent magnet synchronous motor fault diagnosis based on learning," *IEEE Trans. Instrum. Meas*., vol. 68, no. 2, pp. 310-324, Feb. 2019.

[29] X. Ding and Q. He, "Energy-fluctuated multiscale feature learning with deep ConvNet for intelligent spindle bearing fault diagnosis," *IEEE Trans. Instrum. Meas*., vol. 66, no. 8, pp. 1926-1935, Aug. 2017.

[30] J. Feng *et al*., "Injurious or noninjurious defect identification from MFL images in pipeline inspection using convolutional neural network," *IEEE Trans. Instrum. Meas*., vol. 66, no. 7, pp. 1883-1892Jul. 2017.

[31] I. Kamwa *et al*., "Recurrent neural networks for phasor detection and adaptive identification in power system control and protection," *IEEE Trans. Instrum. Meas*., vol. 45, no. 2, pp. 657-663, Jul. 1996.

[32] D. Liu *et al*., "An integrated probabilistic approach to lithium-ion battery remaining useful life estimation," *IEEE Trans. Instrum. Meas*., vol. 64, no. 3, pp. 660-670, Mar. 2015.

[33] I. Nancovska *et al*., "Case study of the predictive models used for stability improvement of the dc voltage reference source," *IEEE Trans. Instrum. Meas*., vol. 47, no. 6, pp. 1487-1491, Dec. 1998.

[34] C. Alippi and V. Piuri, "Experimental neural networks for prediction and identification," *IEEE Trans. Instrum. Meas*., vol. 45, no. 2, pp. 670-676, Apr. 1996.

[35] C. Alippi and V. Piuri, "Neural modeling of dynamic systems with nonmeasurable state variables," *IEEE Trans. Instrum. Meas*., vol. 48, no. 6, pp. 1073-1080, Dec. 1999.

[36] A. H. Tan and K. Godfrey, "Modeling of direction-dependent processes using wiener models and neural networks with nonlinear output error structure," *IEEE Trans. Instrum. Meas*., vol. 53, no. 3, pp. 744-753, Jun. 2004.

[37] A. I. Moustapha and R. R. Selmic, "Wireless sensor network modeling using modified recurrent neural networks: application to fault detection," *IEEE Trans. Instrum. Meas*., vol. 57, no. 5, pp. 981-988, May 2008.

[38] A. Malhi *et al*., "Prognosis of defect propagation based on recurrent neural networks," *IEEE Trans. Instrum. Meas*., vol. 60, no. 3, pp. 703-711, Mar. 2011.

[39] W. Lu *et al*., "Early fault detection approach with deep architectures," *IEEE Trans. Instrum. Meas*., vol. 67, no. 7, pp. 1679-1689, Jul. 2018.

[40] M. Ma *et al*., "Discriminative deep belief networks with ant colony optimization for health status assessment of machine," *IEEE Trans. Instrum. Meas*., vol. 66, no. 12, pp. 3115-3125, Dec. 2017.

[41] B. Taji *et al*., "False alarm reduction in atrial fibrillation detection using deep belief networks," *IEEE Trans. Instrum. Meas*., vol. 67, no. 5, pp. 1124-1131May 2018.

[42] B. Andò *et al*., "Low-order nonlinear finite-impulse response soft sensors for ionic electroactive actuators based on deep learning," *IEEE Trans. Instrum. Meas*., vol. 68, no. 5, pp. 1637-1646, May 2019.

[43] J. Sun *et al.,* "Intelligent bearing fault diagnosis method combining compressed data acquisition and deep learning," *IEEE Trans. Instrum. Meas*., vol. 67, no. 1, pp. 185-195, Jan. 2018.

[44] Z. Chen and W. Li, "Multisensor feature fusion for bearing fault diagnosis using sparse autoencoder and deep belief network," *IEEE Trans. Instrum. Meas*., vol. 66, no. 7, pp. 1693-1702, Jul. 2017.

[45] G. Jiang *et al*., "Stacked multilevel-denoising autoencoders: a new representation learning approach for wind turbine gearbox fault diagnosis," *IEEE Trans. Instrum. Meas*., vol. 66, no. 9, pp. 2391-2402, Sep. 2017.

[46] F. Gu *et al*., "Accurate step length estimation for pedestrian dead reckoning localization using stacked autoencoders," *IEEE Trans. Instrum. Meas*., vol. 68, no. 8, pp. 2705-2713, Aug. 2019.

[47] M. Singha Roy *et al*., "Improving photoplethysmographic measurements under motion artifacts using artificial neural network for personal healthcare," *IEEE Trans. Instrum. Meas*., vol. 67, no. 12, pp. 2820-2829, Dec. 2018.

[48] J. Li *et al*., "CGAN-MBL for reliability assessment with imbalanced transmission gear data," *IEEE Trans. Instrum. Meas*., vol. 68, no. 9, pp. 3173-3183, Sep. 2019.

Mounib Khanafer is an Associate Professor of Electrical and Computer Engineering at the Department of Engineering at the American University of Kuwait, Kuwait. He conducts research in the general areas of Wireless Sensor Networks and Internet of Things and their applications in vehicular networks. He has three years of industrial experience at Nortel Networks, Canada and worked as a Postdoctoral Fellow at the School of Electrical Engineering and Computer Science at the University of Ottawa, Canada from 2012-2013. He received his B.Sc. degree in electrical engineering from Kuwait University, Kuwait in 2002 and his M.Sc. and Ph.D. degrees in electrical and computer engineering from the University of Ottawa, Canada in 2007 and 2012, respectively. He is a Senior Member of IEEE.

Shervin Shirmohammadi (M '04, SM '04, F '17) is currently a Professor with the School of Electrical Engineering and Computer Science, University of Ottawa, Canada, where he is Director of the Distributed and Collaborative Virtual Environment Research Laboratory and his research focuses on multimedia systems and networks, including measurement techniques and applied AI for networking, video streaming, and health systems. He received his Ph.D. degree in electrical engineering from University of Ottawa, Canada and currently serves as the Editor-in-Chief of *IEEE Transactions on Instrumentation and Measurement*.