技術資料

人工知能技術を用いた非破壊計測技術の基礎と最新動向

叶　嘉星*

（2018 年 4 月 16 日受理）

A Survey on Artificial Intelligence Research and Its Applications to Non-destructive Evaluation

Jiaxing YE

Abstract

In the past decade, there has been a revolutionary breakthrough in artificial intelligence research field cultivated by the adoption of deep neural network-based learning architecture for various machine perception tasks. A number of 'human-versus-machine' evaluations demonstrated that state-of-the-art AI-enabled systems are able to substitute the human role in various challenging tasks, such as visual content understanding and speech recognition. This report reviews the major AI concepts pertinent to non-destructive evaluation (NDE) and summarizes over 100 contributions, including both fundamental research and practical applications in the field, most of which appeared in the last five years. Particularly, we survey the use of AI-enabled computing for several critical NDE applications, such as AI approaches to ultrasonic imaging understanding, computerized impact-echo investigation and machine vision approach to displacement measurement of structures. We conclude this survey with a discussion of open challenges together with several inspiring directions for the future research.

Keywords: Artificial Intelligence (AI), non-destructive evaluation, computer vision, ultrasonic imaging, data mining

1. Introduction

All man-made structures, i.e. bridges, dams, and airports, have finite life spans and start to degrade since they are put into service. As the time goes, deteriorations, such as corrosion, fatigue, erosion, wear and overloads, will continue until the structures are no longer fit for their intended use. Maintenance, rehabilitation, and replacement of critical social infrastructures pose worldwide pressing problems to human society. Among all issues related to infrastructure safety management, condition inspection is the foremost one, since it is decision-making stage for any further process, and thus attracted lots of research efforts through decades. In general, damage can be formulated as the change introduced into a system that will deteriorate overall integrity and affect its current or future performance[1]. For instance, a damage in mechanical structures can be defined specifically as change to the material and/or geometric properties. Health condition inspection provides quantitative information on the integrity of the structure. It allows better use of resources than time-based maintenance scheduling, which may be performed even there is no necessity. To achieve on-demand maintenance, a wide variety of condition inspection approaches had been proposed from both theoretical and numerical aspects and those methods can be briefly categorized into two groups: destructive and non-

*分析計測標準研究部門非破壊計測研究グループ

destructive tests[2]. The first group, as the name suggests, would break down the specimen so as to determine its properties, such as strength, toughness and hardness. In contrast, non-destructive evaluation/test (NDE/NDT) refers to the set of non-invasive activities used to assess the condition of objectives or installations without causing any damage. This study mainly addresses the NDE/NDT approaches. Efficient NDE methods, such as the impact-echo test, acoustic emission, ultrasonic inspection and X-ray imaging[3], had been successfully applied in a wide range of applications including detection delamination, inner void of concrete[4]; finding leakage of storage tanks or piping systems[5]; assessment of welding quality and debonding[6]. More recently, the innovative NDE methods become more efficient which enable us get access to health condition of structures, including whether there is damage and what kind of damage it is, and even how severe it is[7].

It is noteworthy that some major drawback existed in the conventional condition inspection methods, i.e. the process is costly and usually finds no faults; also, it is subject to human error in reading resultant data. Advanced sensing technologies and efficient AI-enabled computation modeling are paramount to tackle those shortages. With the rapid progress made in the materials science, it is possible nowadays to design a sensor material that can achieve high sensitivity with smaller size and low-cost. For instance, piezoelectric (PZT) and polyvinylidene fluoride (PVDF) sensors are conventional methods to measure pressure and acceleration of structures which render higher sensitivity compared to the strain gauges[2]. More recently, Micro Electro-mechanical Systems (MEMS) based sensors have become quite active as sensing device which exhibit higher accuracy and lower cost[8]. The booming of new-generation sensing devices leads to a new style of inspection, i.e. the whole process of loading and damaging of target structures can be recorded and such inspection data covers wide timespan and thus is able to offer richer condition information of structures[9]. With the accumulation of long-term multi-modal sensor data, how to efficiently exploit the data becomes challenging issue. The resultant inspection data can be different forms, such as images, videos, sounds and time-series data. The raw data does not tell where/what the flaws are in the structure. Efficient investigation is required to convert raw values into meaningful quantities or semantic judgement that clearly indicate flaws status, i.e. location and severity. Such data interpretation is conventionally performed by skilled experts, i.e. listening to echo sound in hammer sounding test on concrete, examining ultrasonic images for flaw detection of metal/Carbon Fiber Reinforced Polymer Composites (CFRP) structures and reading ground penetration radar images in highway inspection[3,4]. It is evident that human interpretation is high subjective and individual bias can be added to the judgement.

As soon as it was possible to save and load data into a computer, researchers have built systems for automated analysis of non-destructive evaluation data[10,11]. In broad terms, there are two approaches to NDE data analysis: model-driven and data-driven approaches[12]. Model-driven methods establish a high-fidelity physical model of the structure, usually by finite element analysis, and then establish a comparison metric between the model and the measured data from the real structure. If the model is for a system or structure in normal (i.e. undamaged) condition, any departures indicate that the structure has deviated from normal condition and damage is inferred. Data-driven approaches also establish a model, but this is usually a statistical representation of the system, e.g. a probability density function of the normal condition. Departures from normality are then signaled by measured data appearing in regions of very low density[13]. Data-driven methods became quite active in recent years due to the mature applications of low-cost NDE sensor, powerful central processing unit (CPU) and ever-fast high-speed internet access. In nowadays, it is much easier to generate/transfer/store more and more data during structure inspection[2]. To efficiently process the large-scale non-destructive test data, various data-driven analysis algorithms had been developed which are mainly drawn from the discipline of machine learning, or more broadly, artificial intelligence (AI). The design of a data-driven pattern analysis system requires careful attention to the following issues: definition of pattern classes, sensing environment, pattern representation, feature extraction and selection, cluster analysis, classifier design and learning, selection of training and test samples, and performance evaluation. The object of this paper is to illustrate the utility of the data-driven approach to damage identification by means of an extensive survey and by summarizing recent literatures, we aim to:

† present current progress that advanced machine learning (deep learning) techniques have permeated various aspects of non-destructive evaluation research field;

† identify the challenges for successful application of deep learning to structural health condition assessment tasks;

† highlight the potential solutions which may tackle these challenges.

This survey has been structured as follows. In Section 2 we present a concise historical review of artificial intelligence (AI) and machine learning research, and then introduce the representative technical milestones. Particularly, we deliver an overview on Deep Learning – a machine learning paradigm which exhibit strong ability to substitute human's in various visual/acoustic information processing tasks. Section 3 describes the contributions of deep learning to several central tasks in NDE data analysis, including signal enhancement, content-based segmentation, pattern classification, and anomaly detection. Section 4 discusses current open challenges in development of computerized NDE data analysis systems using machine learning techniques, such as lack of data, uncertainty in data annotations. We finalize this study with a summary, an in-depth discussion and an outlook for future works.

2. Review of research progress in Artificial Intelligence

2. 1 AI, the concept

The idea of creating an intelligent machine is as old as computing, if not even older. An initial description had been laid out by Alan Turing in 1950s, who is an English pioneer of computer science. His seminal paper "Computing Machinery and Intelligence" laid out several criteria to assess whether a machine could be said be intelligent, which has since become known as the Turing test[14]. Several years later, the term of Artificial intelligence (AI) had been coined by John McCarthy in 1956 when he organized the famous Dartmouth conference on the subject[15]. Learning, in computer science, is initially defined as a procedure of establishing a model (algorithm) that can perform a specific task, such as visual/acoustic pattern classification[16]. Besides, it has been commonly acknowledged that AI ranges from machines truly capable of thinking to search algorithms used to play board games. The key issues had been extensively studied through decades, including acting and thinking humanly. Figure 1 shows the taxonomy of modern AI research, which comprises of several major subjects including machine learning (ML) and neural networks (NN). Particularly, the active AI technique of deep architecture of learning (DL) is a class of machine learning technique developed largely from 2006[17, 18]. In the following sections, we will present more details of technical and conceptual development in nowadays AI research.

2. 2 Types of machine learning techniques toward AI

Machine learning methods can be generally categorized into two groups: supervised and unsupervised learning, although there are many nuances[19].

In supervised learning, the goal is to learn a mapping from inputs $\boldsymbol{x}$ to outputs y, given a set of data-label pairs $\{\boldsymbol{x},y\}_n$, where $\boldsymbol{x}$, y and n denote the input in the form of vector, the corresponding label and the sample number, respectively. In the following mathematical representation,

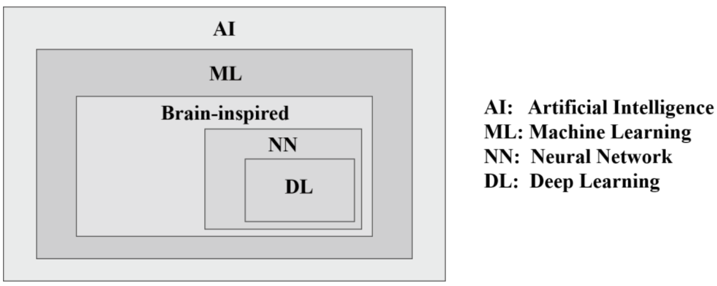

Figure 1 Conceptual coverage map of AI

scalars are noted by lower case letters and vectors are described by boldface lowercase letters. y can be discrete or continuous values, as for the cases of classification and regression. The supervised machine learning scheme commonly amounts to finding model parameters θ that best predict the training data based on an error/loss function $L(y,\tilde{y})$, where $\tilde{y}$ denotes the estimation result of the model obtained by feeding a data point $\boldsymbol{x}$ to the function: $\tilde{y}=f(\boldsymbol{x};\theta)$. The second category, i.e. unsupervised learning algorithms, processes data without labels and are trained to find inner structures of data, such as latent subspaces[13), 18)]. This is sometimes called knowledge discovery[20)]. Unsupervised training can be performed under a batch of loss functions. One example is reconstruction loss $L(\boldsymbol{x},\tilde{\boldsymbol{x}})$ where the model is designated to reconstruct its input, often through a lower dimensional or noisy condition. While, the problem is commonly not well-defined, since we are not told what kinds of structures to capture, and there is no obvious error measurement to use (unlike in the case of supervised learning, where the error is defined as difference between ground truth y to our prediction result for a given x to the observed value). One of the most successful traditional unsupervised learning algorithms are principal component analysis (PCA)[13)] and clustering methods[21)].

In recent decade, several new formulations of machine learning had been proposed, such as reinforcement learning which is grounded on decision theory[16)] to learn how to act or behave when given occasional reward or punishment signals, such as in the famous AlphaGo learned to play go and won top human players[22)]. The techniques become more important in various real applications, including automatic driving and robotics[18)]. This survey mainly focuses on the supervised/unsupervised machine learning, since they are anticipated to play more critical roles in non-destructive test data investigation.

2. 3　AI-enabling techniques, the milestones

Through decades, various machine learning frameworks had been formulated, among which the statistical approach has been most intensively studied and used in practice[23)]. In this section, we review the key contributions that boosted AI research in past half century. We begin the review with a general framework of machine learning system, which includes four steps in a sequential setup. At the first data input stage, AI systems are expected to be able to process the data with various modalities, such as image, audio, and time-series data. Feature extract stage is in charge of generating a vector form from the raw data, which can greatly facilitate further statistical learning with concise representation. At the statistical data analysis step, machine learning techniques are employed to train an optimal model which understands the pattern conveyed by the input data. The model can be regarded as a knowledge base for processing unseen future inputs. Finally, the results are presented with respect to application requirements. The processing flow can be visualized with Figure 2.

From historical viewpoint, AI research can be categorized into four generations, which are shown in the following Table 1. We can see the revolutionary progress had been made in the feature extraction approaches which greatly boost the performance of AI systems to

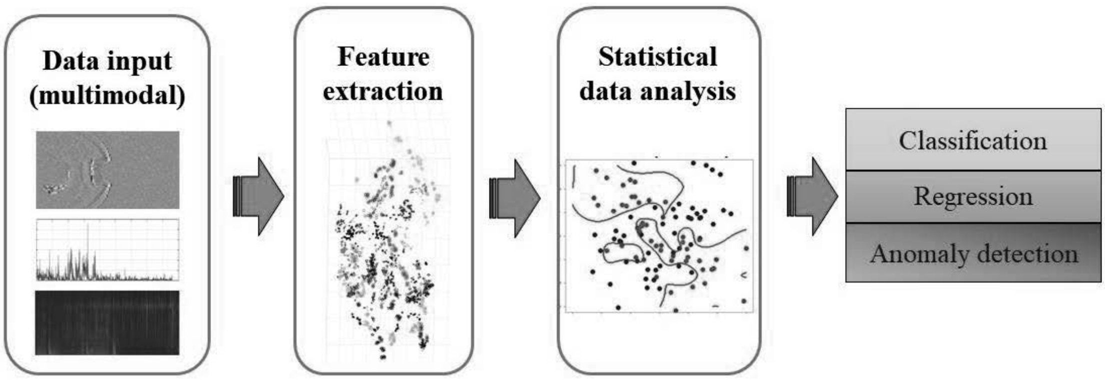

Figure 2　A general processing flow of machine learning system

Table 1　Four generations of AI research

Four generations of machine learning	Feature extraction	Multivariate data analysis
1. Rule based (~1985)	Hand-crafted features (Geometrical)	Correlations Distance measure [16]
2. Traditional Machine Learning (1985~2000)	Hand-crafted features (Geometrical / statistical)	Linear Discriminant function[19]
3. Representation Learning (1998~2010)	Hand-crafted features (Geometrical / advanced statistical)	Non-linear Statistical classifier (SVM[30], RF[35])
4. Deep Learning (2006~)	Automatic hierarchical feature extraction by Deep learning	Softmax [18]/ SVM

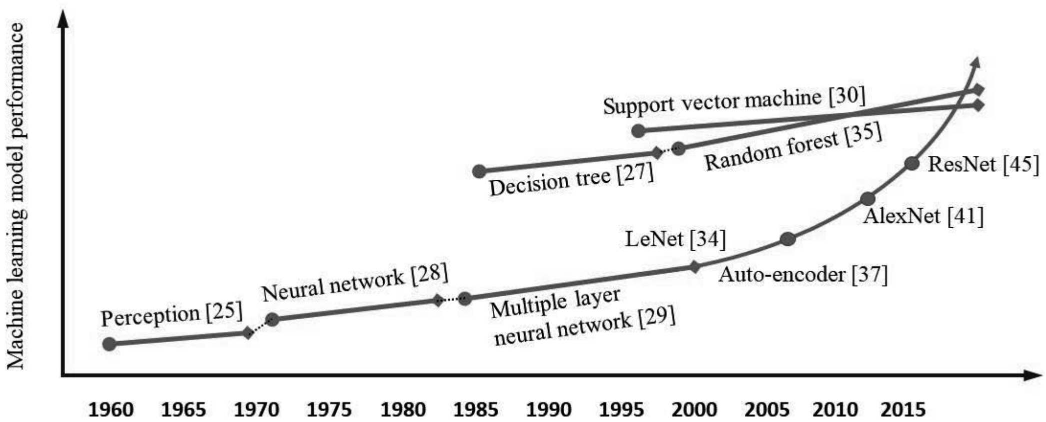

Figure 3　Time line of representative techniques of AI research

tackle real world challenges. Based on the high-level classification of AI paradigms, we present a comprehensive survey on key technical developments.

Figure 3. shows the most remarkable contributions to AI research. According to the figure, the performance of machine learning systems had been constantly increasing through decades and behind this chart, the availability of large scale dataset and high-power computers are also key contributing factors[18]. In parallel to Figure 3. below is a brief historical timeline of machine learning techniques:

1943: McCulloch & Pitts show that neurons can be combined to construct a Turing machine by using ANDs, ORs, & NOTs.[24]. It is the first attempt to build "thinking" machine using computers, meanwhile it gave birth to the research territory of artificial intelligence.

1958: Rosenblatt shows that perceptron algorithm can eventually find a proper classification boundary (convergence) if what they are trying to separate two classes of data which are inherently linear-separable[25]. This work presented the theoretical guarantee of availability of machine learning for easy classification tasks and attracted much attentions from both neural science and computer research fields.

1969: Minsky & Papert showed the computational limitations of Rosenblatt's perceptron with a solid mathematical derivation in the famous book Perceptron[26]. This conclusion was largely responsible for casting serious doubt on the computational capabilities of not only the perceptron, but also neural networks with any forms up to the mid-1980s[18]. During that period, research towards neural network has been regarded as the dead end and almost all research proposals named with "neural network" got killed at grant reviewing.

1984: A special type of classifier is the decision tree[27], which can be trained by an iterative selection of individual features that are most salient at each node of the tree. The criteria for feature selection and tree generation include the information content, the node purity, or Fisher's criterion. Decision tree classification systems had been successfully applied for a number of real applications, such as in[23]. Concurrently, the method is available in the public domain and therefore, often used as a reference benchmark[13), 23)].

1985: The backpropagation algorithm by Geoff Hinton et al[28] revitalizes the field. It is noteworthy that backpropagation and its variants are still being extensively used in current so-called deep learning algorithms and several key parameters, such as learning rate, $1^{st}/2^{nd}$ order optimization schemes, had been investigated with theoretical depth[18].

1988: Neocognitron: a hierarchical neural network capable of visual pattern recognition[29].

1995: One of the most interesting developments in statistical pattern classifier design is the introduction of the support vector classifier by Vapnik[30]. It is primarily a two-class classifier and further generalized to multiple class cases[31), 32)]. The optimization criterion here is the width of the margin between the classes, i.e., the empty area around the decision boundary defined by the distance to the nearest training patterns. These patterns, called support vectors, finally define the classification function. Their number is minimized by maximizing the margin. A batch of distance kernel metrics can be incorporated with support vector machines (SVM) formulation, such as using polynomial and Gaussian radial basis functions. By introducing kernel mapping, SVM offers a possibility to train generalizable, nonlinear classifiers in high-dimensional spaces using a small training set[33].

1998: The most successful type of models for computer vision to date is convolutional neural networks (CNNs) with Backpropagation, which had been proposed by Yan LeCun for hand-written digit recognition[34]. The method achieved record-breaking performance and beat all other methods with large margin. Despite its superiority had been confirmed, the use of CNNs did not gather momentum until related technical requirements have been satisfied, such as efficient training scheme for deep networks, and advances made in computing hardware.

2001: In statistics and machine learning, ensemble methods use multiple learning algorithms to improve the stability and accuracy for better predictive performance. Random forests[35] is one of the most successful ensemble learning method by constructing a multitude of decision trees and outputting the class by majority voting of all trees. It had been validated as efficient approach to tackle the overfitting issues caused by decision tree.

2006: The Hinton lab solves the training problem for Deep Neural Networks for digit handwriting recognition[36), 37)] and opens the door to revolutionary deep learning-based AI research era. The latest progress in Deep learning will be presented with following section.

2. 4　Modern AI — the deep learning era

The swift rise and apparent dominance of deep learning over traditional machine learning methods on a variety of tasks has been clarified in recent years though a batch of human vs. machine evaluations. It is commonly acknowledged that deep learning approaches render three key merits: 1. Universal learning scheme that can deal with data with different forms of data, e.g. video and audio. 2. Deep learning is capable of automatically extract hierarchical features which are robust to the pose/size variations and noises. 3. The learnt model can be efficiently updated as input data arrives[18]. The success of deep learning methods also reflects on the volume of the scientific publications. For instance, deep-learning-related articles in main computer vision venues boosted from fewer than 100 in 2012 to an astounding level of more than 1,000 in 2017. We present a concise review of progress of Deep learning research.

2. 4. 1　Neural networks

The neural networks are defined as one type of learning algorithm that laid fundamentals of current deep learning methods. A neural network is built with a number of neurons or units with some activation a and parameters $\theta = \{\boldsymbol{W}, \boldsymbol{B}\}$, where $\boldsymbol{W}$ and $\boldsymbol{B}$ denote a set of weights and a set of biases, respectively. Notably, both variables are in the matrix form and therefore written in uppercase boldface letters. The activation, which is regarded as core computation process, can be realized by using a linear combination of the input $\boldsymbol{x}$ to the neuron and the parameters, followed by an element-wise nonlinearity σ(·). The process can be referred to as following transfer function:

$$a = \sigma(\boldsymbol{w}^T\boldsymbol{x} + b) \tag{1}$$

There are several options in choosing transfer functions, such as the sigmoid and hyperbolic tangent function. The multi-layered perceptron (MLP), which is one of the most well-recognized structure of neural networks, composes of several stacking-layers of these activations:

$$f(\boldsymbol{x};\Theta) = \sigma(\boldsymbol{W}^T\sigma(\boldsymbol{W}^T...\sigma(\boldsymbol{W}^T\boldsymbol{x} + b)) + b). \tag{2}$$

Here, $\boldsymbol{W}$ is a matrix is composed of columns $\boldsymbol{W_k}$, associated with activation k in the output. Between the data input and prediction output, we can see several layers and they are often referred to as hidden layers. As for the case that a neural network contains multiple hidden layers (i.e. more than 3), it is typically called a deep neural network, hence leading to the term deep learning. At the final layer of the network the activations are mapped to a posterior distribution over all classes memberships $P(y|\mathbf{x}:\theta)$ through a softmax function[27]:

$$P(y|\boldsymbol{x};\Theta) = \mathrm{softmax}(\boldsymbol{x};\Theta) = \frac{e^{w_k^T x + b}}{\sum_{k=1}^{K} e^{w_k^T x + b_k}}, \tag{3}$$

where $\mathbf{w}_k$ denotes the weighting vector leading to the output node associated with class k. A systematic diagram of three-layer MLP is shown in Figure 4.

One critical issue is how to find the best setting of parameters θ of a neural network dedicated for a given task, such as recognizing objects in an image or classifying specific sound from environment audio data.

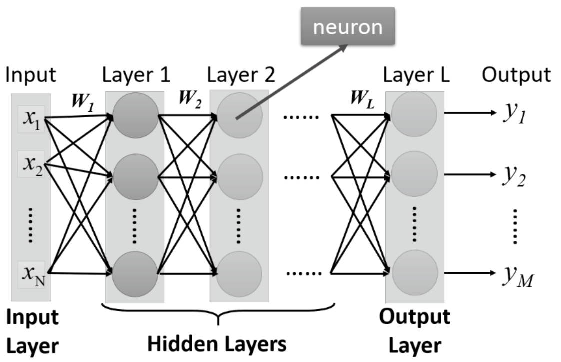

Figure 4 Diagram of typical multiple layer neural network architecture

Concretely, it can be achieved by numerical optimization. A cost function J(θ) is initially defined which typically includes a performance measure evaluated on the entire training set as well as additional regularization terms[19), 20)]. Optimization, simply speaking, is to find a rule to significantly reduce cost function J(θ) through updating all the parameters in the neural network. Through decade, stochastic gradient descent (SGD) has been the most popular method to fit parameters θ to a dataset[34]. In SGD, a small subset of the data, i.e. commonly named as mini batch, is employed for each gradient update instead of the full data set. Optimization, i.e. achieving the maximum likelihood, in practice amounts to minimizing the negative log-likelihood:

$$\arg\min_{\theta} - \sum_{(n=1)}^{N} \log[P(y_n|\mathbf{x}_n;\theta)]. \tag{5}$$

For a long period, deep neural networks (DNN) with multiple hidden layers were considered difficult to train efficiently. They only gained popularity in 2006[17] when it was shown that training DNNs layer-by-layer in an unsupervised manner (pre-training), followed by fine-tuning of the stacked network in a supervised manner, could achieve good performance. Two popular architectures trained in such a way are deep belief networks (DBNs)[36] and stacked auto-encoders (SAEs)[37]. However, these techniques are rather sophisticated and require a batch of engineering tricks to generate promising results. Currently, the most popular scheme is to train models in an end-to-end fashion, and effectively simplifying the training process. The most popular learning schemes are convolutional neural networks (CNNs)[38] and recurrent neural networks (RNNs)[18]. CNNs are currently most widely used in (medical) image analysis, although RNNs are gaining popularity in time-series analysis. The following sections will give a brief overview of each of these methods, starting with the most popular ones, and discussing their differences and limits when applied to NDE data investigation.

2. 4. 2 Convolutional Neural networks (CNN)

CNN denotes a family of neural network architecture which are dedicated to process matrix-shaped data, i.e. images. Briefly, there are two major changes between MLPs and CNNs. First, the weights in CNNs are shared a manner of performing convolution operations on 2-dminensional data[39]. This way, the model does not need to establish individual detectors for same kind of object presenting at different locations in an image, enabling the network equivariant with respect to positioning-translations of the input. It also efficiently reduces the total amount of parameters (i.e. the number of weights no longer depends on the size of the input image) that need to be learned. An example of a 1D CNN is shown in Figure 5.

From the chart, we can see several layers are stacked in a sequential manner; at each layer, the input image is convolved with a set of K kernels $\boldsymbol{W} = \{\boldsymbol{W}_1, \boldsymbol{W}_2, \cdots, \boldsymbol{W}_K\}$ and added biases $\boldsymbol{b} = \{b_1, b_2, \cdots, b_k\}$, each generating a new feature map $\boldsymbol{X}_k$. These features are subjected to an elementwise non-linear transform $\sigma(\cdot)$ and the same process is repeated for every convolutional layer l:

$$\boldsymbol{X}_k^{l} = \boldsymbol{X}_k^{l-1} * \boldsymbol{X}^{l-1} + \boldsymbol{b}_k^{l-1} \tag{6}$$

The second core difference between CNNs and MLPs lies in the adoption of pooling layers in CNNs, where pixel values of neighborhoods are aggregated using a permutation invariant function, typically the max or mean operation. Such process can induce favorable translation invariance to the feature maps and further eliminates redundant parameters in the network as well. Fully-connected layers (i.e. regular neural network layers) are often added to the final stage of the network, where weights are no longer shared. Similar to MLPs, a distribution over possible classes is generated by feeding the activations in the final layer through a softmax function and the network is trained using maximum likelihood.

2. 4. 3　Deep CNN Architectures

Given the prevalence of CNNs in multimodal data analysis, we elaborate on the most widely-applied architectures and their differences. LeNet[34], Auto Encoder[40] and AlexNet[41], introduced over a decade later, were in essence very similar, e.g. both schemes were relatively shallow, consisting of two and five convolutional layers, respectively; employed kernels with large receptive fields in layers close to the input and smaller kernels closer to the output. Particularly, AlexNet did incorporate rectified linear units (ReLU) as activation function that is different from conventional hyperbolic tangent setting.

After 2012 the exploration of novel architectures took off, and in the past five years a trend emerged to build far deeper and wider models. By stacking smaller kernels, it is possible to represent a function by using fewer parameters. Meanwhile, such deeper architectures generally render a lower memory footprint through statistical inference, which enable their deployment on smartphones and other mobile computing devices. A notable article performed in-depth investigation on much deeper networks by employed smaller, fixed size kernels in each layer[42]. The most remarkable work is a 19-layer model often referred to as VGG19 or OxfordNet won the ImageNet challenge of 2014.

On top of the deeper networks, more complex designs have been exploited aiming at improving model training efficiency and again reducing the amount of parameters. A 22-layer network named GoogLeNet was introduced in[43], also referred to as Inception, which made use of a couple of inception blocks, a module with a set of

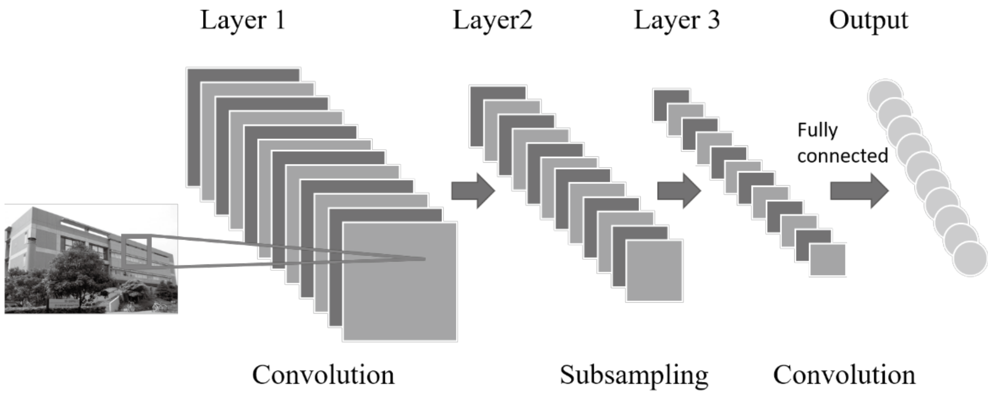

Figure 5　General framework of deep neural network for computer vision

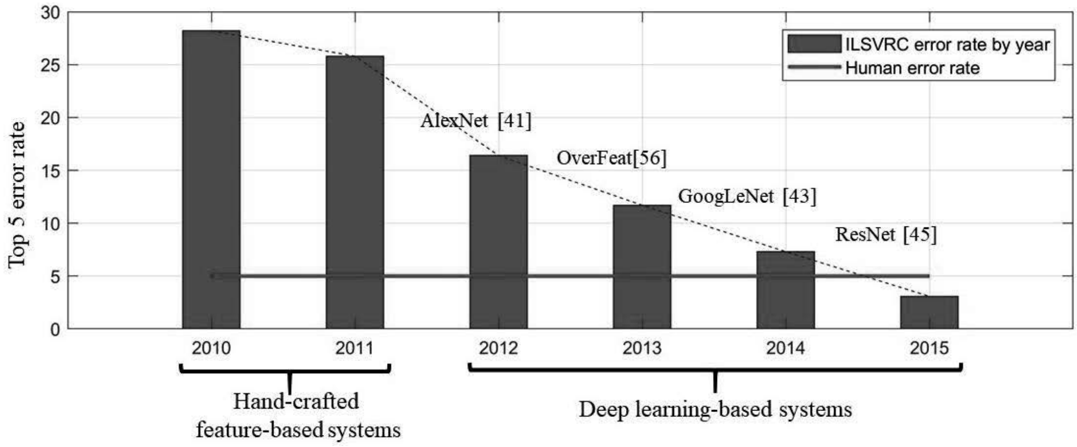

Figure 6 Advancement of machine vision for object recognition on the ILSVRC challenge. Since deep learning had been introduced in 2012, the error rates continued to decrease, and the latest model outperformed human ability in 2015

convolutions of different sizes. Network in network (NIN) is another representative network design that can extract rich information of raw data by introducing deeper net structure[44]. It is noteworthy that the ResNet architecture won the ImageNet challenge in 2015[45] and consisted of so-called ResNet-blocks. Rather than inferring a function, the residual block only estimates the residual and is thereby biased towards learning mappings of each layer. Such design will enhance reconstruction property of feature mapping. The impressive progress of computer vision systems for object recognition has been boosted by deep learning schemes and we summarize the process in Figure 6. This way, even deeper models can be trained effectively. Notably, only the results until 2015 had been shown because since 2015, the performance on the ImageNet benchmark has saturated and it is difficult to assess whether the small increases in performance can really be attributed to better and more sophisticated architectures. The advantage of the lower memory footprint these models provide is typically not as important for real applications. Consequently, AlexNet or other simple models such as VGG are still popular for non-destructive evaluation data processing, though recent advanced results are preferred to use a version of GoogleNet called Inception version3[46), 47)]. There are two major reasons accounted for such status, which are superiority of learning architecture and public availability of open source codes.

2. 4. 4 Recurrent Neural Networks (RNNs)

RNNs were well-developed tools for discrete sequence data analysis in deep learning field[20]. They can be understood as a variant of MLPs with both the input and output can be of varying length. One representative application of RNNs is machine translation where a sentence of the source and target language are the input and output[18]. As for classification tasks, the model infers a distribution over classes $P(y|\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_T;\theta)$ given a sequential observation $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_T$, rather than a single input vector $\boldsymbol{x}$. The plain RNN maintains a latent or hidden state h at time t that is the output of a non-linear mapping from its input $\boldsymbol{x}_t$ and the previous state h_{t-1}:

$$\boldsymbol{h}_t = \sigma(\boldsymbol{W}\boldsymbol{x}_t + \boldsymbol{R}\boldsymbol{h}_{t-1} + \boldsymbol{b}) \tag{7}$$

where weighting matrices $\boldsymbol{R}$ and $\boldsymbol{W}$ are shared over time. For classification, one or more fully connected layers are typically added followed by a softmax function to map the sequence to a posterior over the classes.

$$P(y|\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_T;\Theta) = \mathrm{softmax}(\boldsymbol{h}_t;\boldsymbol{W}_{out}, \boldsymbol{b}_{out}) \tag{8}$$

Since the gradient needs to be backpropagated from the output through time, RNNs are inherently deep (in time) and consequently suffer from the same problems with training as regular deep neural networks[48]. To this end, several specialized memory units have been developed, the initial work is the Long Short Term Memory (LSTM) cell[49], which is still popular concurrently.

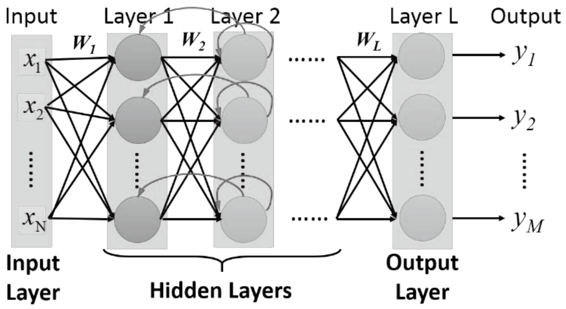

Figure 7 A typical RNN neural network, please note the red arrow lines, which allow recurrent jumps of information passing

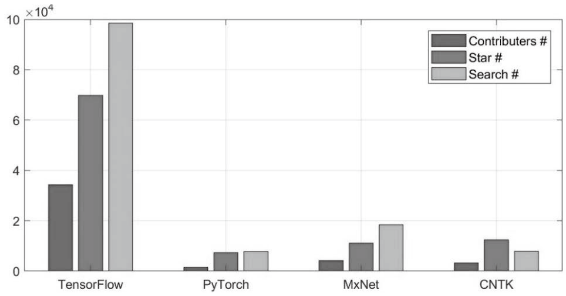

Figure 8 Popularity comparison between latest open-source deep learning libraries. The data is collected from GitHub (https://github.com), a world's leading software development platform. A large-margin lead can be seen of TensorFlow in terms of all metrics.

One remarkable work towards development of RNNs is the Gated Recurrent Unit[50], which can be regarded as a recent simplification of the LSTM. Although initially designated for one-dimensional data, RNNs are increasingly employed to process images. For instance, so-called pixelRNNs are used as autoregressive models, generative models that can eventually produce new images similar to samples in the training set[51].

2. 4. 5 Unsupervised deep learning

Besides the supervised machine, in which training data is provided together with ground truth labels, deep learning can also be formulated to tackle label-free data, such as in dimension reduction[12] and data-driven anomaly detection[13]. The popular techniques include Auto-encoder (AE) and Restricted Boltzmann Machines (RBMs) and Deep Belief Networks (DBNs)[17, 18]. The detail math derivations can be found in the cited works. However, this review is mainly focused on supervised learning paradigm because it the main solution applied for non-destructive test data analysis.

2. 4. 6 Hardware and Software for deep learning

One of the main driving force to steep rise of deep learning has been the widespread availability of Graphics processing unit (GPU) and GPU-computing libraries (CUDA, OpenCL). GPUs are highly parallel computing engines, which have an order of magnitude more execution threads than central processing units (CPUs). With current hardware, deep learning on GPUs is typically 10 to 30 times faster than on CPUs. More recently, field programmable gate arrays (FPGAs) have been adopted for accelerating the implementation of deep learning networks due to their ability to maximize parallelism as well as due to their energy efficiency[52]. Next to hardware, the other contributor behind the popularity of deep learning methods are the open source software packages. These libraries provide efficient GPU implementations of important operations in neural networks, such as convolutions; allowing the user to implement ideas at a high level rather than worrying about low-level efficient implementations. At the time of writing, the most popular packages were (in alphabetical order):

・Caffe[52]. Provides C++ and Python interfaces, developed by graduate students at UC Berkeley.

・CNTK[53]. Named as Microsoft Computational Network Toolkit which provides an open-source toolkit for commercial-grade distributed deep learning. It was developed by Microsoft.

・Tensorflow[54]. Provides C++ and Python and interfaces, developed by Google and is used by Google research.

・PyTorch[55]. Provides a Lua interface and is used by, among others, Facebook AI research.

・MxNet[56]. Provides a very flexible machine learning library which is computation and memory efficient and supports various platforms ranging from mobile devices to distributed GPU clusters. The library was developed by graduate students from Carnegie Mellon University and the library is currently used by Amazon. co.

There are third-party packages written on top of one or more of these frameworks, such as Lasagne (https://github.com/Lasagne/Lasagne) or Keras (https://keras.io/). Those high-level application programming interfaces

(API) efficiently facilitate fast development of AI-enabled applications, however the efficiency was not satisfying. By comparison, Tensorflow and CNTK are well-suited for practical applications, i.e. building product recommendation engine for online shopping and web-scale image retrieval. While, it has been criticized to be complex to use and efficiency is not high (single computer-wise). PyTorch is more welcomed by academic committee due to the high flexibility in deep learning structure configuration. MxNet renders user-friendly language interface and a high proficiency in multiple-GPU acceleration.

3. AI applications in NDE fields

The last five years have seen remarkable progress in machine learning research, and a spreading trend emerged to develop human-level machine learning systems to relieve people from laborious and exhausting tasks in the non-destructive test for infrastructures. In order to substituting human role in hammering response interpretation and to achieve optimal decision-making in NDT, great efforts had been carried out to establish data-driven machine learning system to understand NDE data in terms of various forms, i.e. image, video, and time-series signal[2), 14)]. We present a review on current research status as follows.

3. 1 Emerging topic of computerized ultrasonic imaging analysis

Ultrasonic imaging inspection systems had been extensively applied for NDT due to several favorable merits, such as high sensitivity to most material damage, and proficiency in extraction of defect location and size specifications[57)]. The principle of ultrasonic testing is based on detection and analysis of received ultrasonic waves, from which defect-induced patterns can be clearly observed. A typical ultrasonic imaging inspection system is composed of three parts: A laser scan unit mounted the computer controlled mechanical stage which generates the excitation of ultrasound due to thermal expansion. A transducer attached to the surface of the specimen collects the ultrasonic waves propagated through the specimen. Through an amplifier and a digital oscilloscope (A/D converter), the received signals are transmitted to a computer and stored in the computer hard drive. In current applications, ultrasonic image data requires an inspection engineer to determine if there are any defects present. The inspection performance, therefore, is to a large extent depending on inspector's technical skill and the assessment results may vary considerably due to human factors. Over the past decade, there has been a dramatic increasing interest in research towards automated assessment for ultrasonic inspection using machine learning technique[58)–78)]. Figure 9 presents a diagram describing both the hardware setup of ultrasonic imaging inspection system and human/computerized

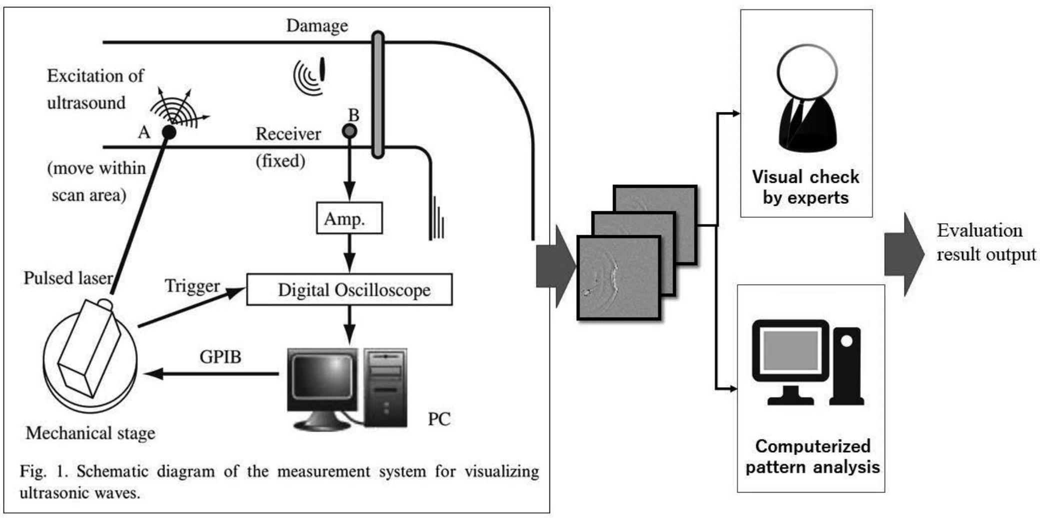

Figure 9　Diagram of ultrasonic scanning system for non-destructive test with human/machine data interpretation

defective pattern investigation schemes.

Plenty of research efforts had been delivered to design efficient automatic ultrasonic data interpretation systems, i.e. ultrasonic echo waveform and wave images had been investigated by using advanced machine learning classifiers. We present the comprehensive review using following table, in which the ultrasonic signal representation and statistical machine learning algorithms were summarized with two columns.

According to the above table, we come to several conclusions: first, a significant trend can be seen that AI-enabled ultrasonic inspection data interpretation emerged as active research topic in recent 5 years and a wide variety of advanced machine learning techniques had

Table 2　Overview of papers using AI techniques for ultrasonic non-destructive tests.

Input signal	Year	Feature representation	Multivariate data analysis
Ultrasonic waveform	58), 2002	Wavelet transform	Shallow artificial neural networks with RBF activation function
	64), 2009	Ultrasonic echoes with matching pursuit (MP) features	Sparse deconvolution method for signal enhancement
	66), 2012	Ultrasonic echo waveform	Sparsity-induced feature learning for noise reduction
	59), 1996	Ultrasonic A-scan signal with principal component analysis	Shallow artificial neural networks
	67), 2013	Ultrasonic A-scan signal	Self-organizing Maps classifiers
	68), 2015	Ultrasonic A-scan signal	Support Vector Machines (SVM)
	70), 2016	Ultrasonic A-scan signal with wavelet transform	Hidden Markov Model (HMM)
	71), 2017	Vibrothermography images	Maximum likelihood estimates (MLEs)
	72), 2017	Ultrasonic A-scan signal	Convolutional Neural Networks (CNNs)
	74), 2017	Ultrasonic A-scan signal	Split Spectrum Processing (SSP) and (shallow) artificial neural networks
	75), 2017	Ultrasonic A-scan signal Fourier spectrum with band feature selection	2-layer perceptron (MLP) neural network
	77), 2017	Ultrasonic A-scan signal	Dictionary learning using K-SVD
Ultrasonic image	60), 1996	Time-of-flight diffraction (TOFD) scan image	Shallow artificial neural networks
	61), 1997	Ultrasonic B-scan image	Image histograms thresholding scheme
	62), 2006	Phase information extracted from TOFD images	Cross-correlation coefficient
	63), 2007	Co-occurrence based matrix features	Multilayer neural-fuzzy network
	65), 2011	Ultrasonic B-scan image with time-frequency analysis	Shallow artificial neural networks
	69), 2016	Ultrasonic B-scan image	Sparse deconvolution method
	73), 2017	Ultrasonic B-scan image	Convolutional Deep Belief Networks (CDBN)
	76), 2017	Ultrasonic B-scan image with Hilbert–Huang transform (HHT)	Cross-Correlation
	78), 2017	Ultrasonic guided wave images	4-layer deep neural network

been evaluated for the task[68)-78)]. Second, as for the input ultrasonic signal, there are two groups: the raw echo signal and ultrasonic B-scan images. Accordingly, different feature extraction approached had been employed to deal with the data, i.e. Fourier and wavelet transform had been introduced to process A-scan waves and co-occurrence features had been investigated for ultrasound B-scan images as well. Third, advanced statistical pattern analysis approaches, such as sparse coding[69),77)] and support vector machines[68)], had been evaluated for the ultrasonic signal analysis. Those literatures confirmed that novel machine learning/pattern recognition techniques can substantially contribute to ultrasonic data interpretation. It is noteworthy that as shown in the most recent publications, deep neural networks had been repeatedly mentioned[72),73)], while the most complex neural network up to now applied for the task was limited to 4 layers[78)], which is genuinely NOT a deep learning solution[78)]. The major factor accounted for this is that the ultrasonic inspection datasets are confined to small-scale so that there is no significant benefit to employ deep neural networks for generating efficient feature representations. It is also unfortunate that there is no standard/public ultrasonic inspection database and thus, each researcher has to generate their own. Hence, it is not possible to come to a conclusion and handpick the best.

According to the above survey, we also discovered several possible directions to conduct further research. In the first place, we found that current literatures usually dealt with ultrasonic signal in the form of waveform and 2-dimensional images, while the short-time dependences were ignored. In Figure. 10, we present a chart to describe the current status, in which A-scan waveform and B-scan images of ultrasonic signal had been shown in the left and middle, respectively. To our interest, the rich discriminant information of ultrasonic signal is conveyed in spatio-temporal formulation, which can be seen in the rightmost plot in Figure 10. In such setting, we regard the ultrasonic wave propagation image sequences as one video clip and between-frame information herein will be characterized for the defect-induced ultrasonic wave pattern investigation. Meanwhile, latest research progresses in machine learning and computer vision field have deemed that deep learning can greatly facilitate video analysis with automatic feature engineering[79)-81)]. With the massive data collection, it can be anticipated that deep learning-based learning schemes for ultrasonic video frames analysis will outperform the conventional waveform/image-based approaches.

3. 2 Emerging topic of AI-enabled impact-echo test

The impact-echo, due to its cost efficiency and simplicity, has been extensively applied for concrete structure condition assessment over decades[82)]. It played a key role as efficient Non-destructive test (NDT) method to detect multiple defects of concrete, e.g. delamination and inner voids[83)]. In general, the method consists of several steps: 1. a hammer impact is applied to surface of concrete structure in order to generate elastic stress waves. 2. stress waves propagate inside the structure and then transmit through the air. 3. inspection workers will examine the echo signal and determine health condition of concrete. Although the technique is widely applied all over the world, there remains several inherent drawbacks, such as erroneous data interpretation due to subjective

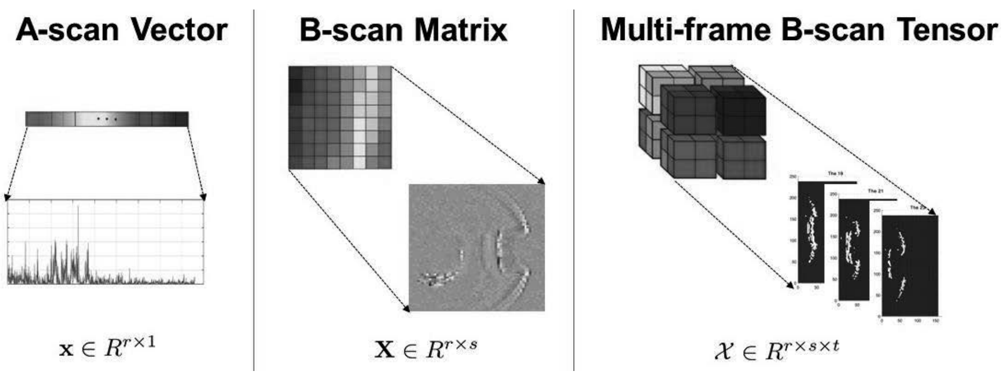

Figure 10 Three formulations of ultrasonic signal input for defect detection

judgement and requiring expertise knowledge. To tackle the limitations, research towards developing AI-enabled automatic echo analysis system for structural health assessment garnered a lot of attentions in recent years[84), 85)]. Signal processing and machine learning techniques are integral components to build human-like echo analysis engine. In Figure 11 we show a basic diagram of computerized echo analysis system for air-coupled impact-echo test. Through decades, plenty of theoretical and empirical studies on impact-echo methods had been carried out and major results have been reported. In[86)], by adopting wavelet Transform (WT), echo waveform is converted to frequency domain and spectral analysis is performed subsequently. Extensive studies revealed that there exists an empirical function describing relationship between peak frequency in echo amplitude spectrum and depth of inside defect, which can be expressed as follows:

$$d = \beta \frac{C_p}{2f_{\mathrm{peak}}} \tag{9}$$

where f_{peak} denotes peak frequency of echo signal spectrum, C_p is the velocity of the longitudinal, β is constant of 0.96 for plate-shape structures wave[83)] and d represents depth of inside void. However, some recent studies reveal the availability of formula (9) is constrained by the size and flatness of defect area, e.g. it is only valid for the case that void is parallel to surface, otherwise the echo resonance will behave differently from Eq. 9. Such limitation opens doorway to data-driven statistical pattern analysis of echo signals for defect detection.

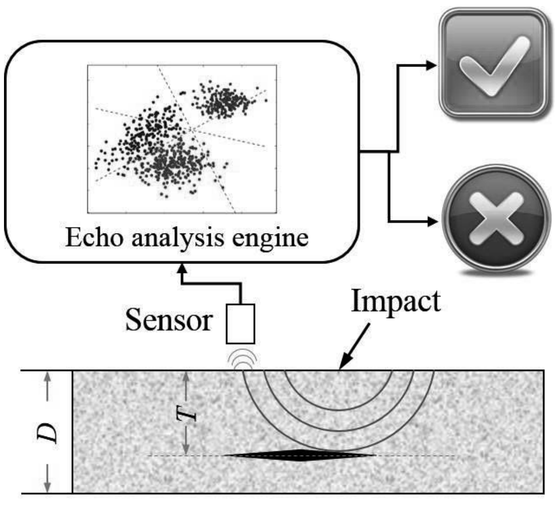

Figure 11　Diagram of AI-enabled impact-echo system

The initial systems commonly deal with echo investigation problem using statistical pattern classification, in which various conventional classifiers have been employed to classify echo spectra obtained from healthy/defective concretes. The representative machine learning models for impact-echo are Gaussian mixture models (GMM)[87)], Artificial Neural Network (ANN)[87)] and Support Vector Machines (SVM)[84)], to characterize discriminant information of healthy/defective echoes. In recent years, significant progress has been made in noise robust echo feature representation learning. Advanced signal descriptors developed by the bag-of-words model (BoW model)[88)] and sparse coding approaches[89)] has been proved to be effective for anomalous echo identification under hostile acoustic environment. It is noteworthy that these literatures commonly assume that all training and test hammer responses are sampled from same population; the experimental dataset was confined to be the laboratory-scale as well. It is anticipated to be problematic when we directly apply the analysis model trained by lab-scale data to practical hammer sounding test, because the pre-collected training data is quite limited to render sufficient discriminant information to deal with complex real echoes. How to deal with incremental data of impact-echo test can be one interesting theme which has both research significance and application impact.

4. AI-enabled NDE applications: open challenges and potential solutions

Based on the above-presented survey, we outline several common issues in AI-enabled NDE system development together with several possible research directions to leverage the issues.

4. 1　Computerized non-destructive test data analysis with upgradable design

Current machine learning schemes applied for the non-destructive evaluation are restricted to the standard batch setting, which assumes that both training and testing data reside in the same feature space with the static statistical characteristic; hence, model training can be performed over the pre-collected laboratory-scale database[60),65),67),68),72),73),75),78),84),88),89)]. In practice, however, such an assumption does not hold. The patterns of onsite NDE data can alter significantly with the specifications of testing objectives, such as material, shape and years of service[83)]. From the viewpoint of statistical learning, these factors would make the posterior distribution of the test data drift from that of the pre-collected training samples; thus, degrading the healthy/defective pattern discrimination performance. In real application scenario, it is indispensable to adopt an alternative hypothesis which inherently admits that the pre-collected training dataset only covers small range of the whole real-world data distribution. To this end, a new formulation of NDE data pattern classification with the online learning paradigm is essential, in which efficient model updating schemes have been exploited to minimize the cumulative prediction loss suffered along with the continuous input of data. Online learning is a well-established learning scheme which has both theoretical and practical appeals[90),91)] and it is particularly well-suited to the non-destructive test data investigation, since the large-scale response data can be accessed only in a sequential way.

4. 2 Efficient pattern characterization from limit data with annotations

Fueled by several factors, the AI technologies already pervade our lives[92)]. One of the foremost factors is the large-scale data collection, such as the ImageNet dataset utilized for computer vision research consists of 14 million images from 21841 classes[41)]. However, as for the application of non-destructive test, huge amount of data collection with expert labels are usually infeasible due to the high-cost of time and budget. Through this survey, we had been looking at current trends in the machine learning research and searching for key areas that could indeed leverage the limit in data capture.

4. 3 Semi-supervised learning scheme.

Conventional machine learning systems developed for non-destructive test mainly adopted supervised learning scheme, in which a completely labelled dataset is provided in advance and statistical learning is highly tailored to particular tasks, i.e. impact-echo[89)] and ultrasonic data interpretation[77)]. However, manual labelling often takes considerable efforts from skilled human agent and thus it can be unaffordable to generate annotations to all the instances. Latest machine-learning research have proved that unlabeled data, when used in conjunction with a small amount of labeled data, can produce superior results in pattern investigation accuracy[93),94)]. Since acquisition of unlabeled data is relatively inexpensive compared to the fully labelled data collection, semi-supervised learning can be of remarkable practical value. Another reason to assume that semi-supervised methods will possess a significant role to play is the analogue to human learning[95)], which seems to be much more data efficient; we can learn to recognize objects and structures without knowing the all the labels. Instead, we only need very limited supervision at the beginning for the task. Semi-supervised learning is also of theoretical interest in deep learning context[96)].

4. 4 Data augmentation using GANs.

Another way to tackle lack of data issue is to generate fake data which is expected to be helpful to train efficient model for NDE data pattern investigation. There are two novel strategies which could have an impact: variational auto-encoders (VAEs), introduced by Kingma and Welling[97)] and generative adversarial networks (GANs), introduced by Goodfellow et al.[98)]. The former embedded variational Bayesian graphical models into neural networks as encoders/decoders. The latter employs two competing convolutional neural networks where one is to generate artificial data instances and the other is discriminating fake from real samples. Both networks are generative networks with stochastic components. Most importantly, they can be trained in an end-to-end fashion and the features can be learnt without supervision[99)]. As mentioned in previous paragraphs, unlabeled data is much easier to collect can therefore VAEs and GANs could optimally leverage this wealth of information. One successful application for aircraft engines failure predication had been reported in[100)].

5. Conclusion

Machine learning algorithms, in particular the latest convolutional networks with deep-stacking structures, have emerged as a solid selection for non-destructive evaluation data analysis. This study reviewed the latest progress of machine learning techniques introduced to the non-destructive evaluation field, most of which have been released within the last 5 years. We survey the advancement of machine learning techniques from both theory and real application perspectives. Also, we examined the published literatures on the use of machine learning for two applications of our interest: ultrasonic inspection data investigation and computerized impact-echo test. Based on concise overviews, we outlined common challenges for development of AI-enabled non-destructive test systems. Furthermore, we pointed out several inspiring research directions to tackle those challenges.

6. Reference

1) Doyle, M. W., Stanley, E. H., Havlick, D. G., Kaiser, M. J., Steinbach, G., Graf, W. L. and Riggsbee, J. A.: Aging Infrastructure and Ecosystem Restoration. Science, 319-5861 (2008) 286-287.

2) Christian Boller, Fu-Kuo Chang and Yozo Fujino: Encyclopedia of Structural Health Monitoring, (John Wiley & Sons, New York, 2009) Chapter 3-6

3) Sohn, H., Farrar, C. R., Hemez, F. M., & Czarnecki, J. J.: A Review of Structural Health Review of Structural Health Monitoring Literature 1996-2001, (Los Alamos National Laboratory, 2002).

4) Ciolko, A. T., and H. Tabatabai.: Nondestructive methods for condition evaluation of prestressing steel strands in concrete bridges, (NCHRP Project Final Report, 1999) 10-53.

5) Chang, Peter C., Alison Flatau, and S. C. Liu: Health monitoring of civil infrastructure, Structural health monitoring 2-3 (2003) 257-267.

6) Hayashi T, Rose JL.: Guided Wave Simulation and Visualization by a Semianalytical Finite Element Method. Materials Evaluation. Jan (2003) 75-79.

7) Doebling SW, Farrar C, Prime MB, Daniel WS: Damage identification and health monitoring of structural and mechanical systems from changes in their vibration characteristics: a literature review. (LA-13070-MS, May 1996)

8) Carden, E. Peter, and Paul Fanning: Vibration based condition monitoring: a review. Structural health monitoring 3-4 (2004) 355-377.

9) Nagayama, Tomonori, and Billie F. Spencer Jr.: Structural health monitoring using smart sensors. (Newmark Structural Engineering Laboratory. University of Illinois at Urbana-Champaign., 2007)

10) Alander JT: Population size, building blocks, fitness landscape and genetic algorithm search efficiency in combinatorial optimization: an empirical study. (Chambers LD (eds) Practical handbook of genetic algorithms: complex coding systems, CRC Press, New York 1999) Chapter 13

11) Xu YG, Liu GR, Wum ZP, Huang XM: Adaptive multilayer perceptron network for detection of cracks in anisotropic laminated plates. Int J Solids Struct 38 (2001) 5625-5645

12) Gopalakrishnan, Srinivasan, Massimo Ruzzene, and Sathyanaraya Hanagud.: Computational techniques for structural health monitoring. (Springer Science & Business Media, 2011)

13) Christopher M. Bishop: Pattern Recognition and Machine Learning (Information Science and Statistics), (Springer-Verlag New York, Inc., 2006)

14) Turing, Alan M.: Computing machinery and intelligence: Parsing the Turing Test. (Springer, Dordrecht, 2009) 23-65.

15) John McCarthy: Programs with Common Sense. Conf. (Teddington Conference on the Mechanization of Thought Processes, 1958)

16) Russell S J, Norvig P.: Artificial intelligence: a modern approach (Pearson Education Limited, Malaysia, 2016)

17) Hinton, G. E., Salakhutdinov, R. R.: Reducing the dimensionality of data with neural networks. Science 313 (2006) 504-507.

18) Goodfellow, I., Bengio, Y., Courville, A: Deep Learning, Cambridge: MIT press (2016).

19) Duda, R. O., Hart, P. E., & Stork, D. G.: Pattern classification. 2nd. (Edition. New York, 55, 2001)

20) Kevin Murphy: Machine Learning-A Probabilistic Perspective, Cambridge: MIT press (2016).

21) Fraley, Chris, and Adrian E. Raftery.: How many clusters? Which clustering method? Answers via model-based cluster analysis. The computer journal, 41-8 (1998) 578-588.

22) David Silver, Julian Schrittwieser, Karen Simonyan, Ioannis Antonoglou, Aja Huang, Arthur Guez, Thomas

Hubert, Lucas Baker, Matthew Lai, Adrian Bolton, Yutian Chen, Timothy Lillicrap, Fan Hui, Laurent Sifre, George van den Driessche, Thore Graepel & Demis Hassabis: Mastering the game of Go without human knowledge, Nature, 550 (2017) 354−359

23）Jain, A. K., Duin, R. P. W., & Mao, J.: Statistical pattern recognition: A review. *IEEE Transactions on pattern analysis and machine intelligence*, 22−1(2000), 4−37.

24）McCulloch, Warren S., and Walter Pitts: A logical calculus of the ideas immanent in nervous activity, The bulletin of mathematical biophysics 5−4 (1943): 115−133.

25）Rosenblatt, Frank: The perceptron: A probabilistic model for information storage and organization in the brain, Psychological review 65−6 (1958) 386.

26）Minsky, Marvin L.: c*omputation: finite and infinite machines*. Prentice-Hall, Inc., 1967.

27）L. Breiman, J.H. Friedman, R.A. Olshen, and C.J. Stone: Classification and Regression Trees. (Wadsworth, Calif., 1984)

28）Ackley, David H., Geoffrey E. Hinton, and Terrence J. Sejnowski. A learning algorithm for Boltzmann machines. Cognitive science 9−1 (1985): 147−169.

29）Fukushima, Kunihiko. Neocognitron: A hierarchical neural network capable of visual pattern recognition. Neural networks 1−2(1988)119−130.

30）Cortest, Corinna, Vapnik, Vladimr N. Support Vector Networks, Machine Learning, 20−3 (1995) 273−297

31）V.N. Vapnik, Statistical Learning Theory. (New York: John Wiley & Sons, 1998)

32）B. Scho "lkopf: Support Vector Learning, (Ph.D. thesis, Technische Universita" t, Berlin, 1997)

33）Shawe-Taylor, John, and Nello Cristianini: Kernel methods for pattern analysis. (Cambridge university press, 2004)

34）LeCun, Yann, et al.: Gradient-based learning applied to document recognition, Proceedings of the IEEE 86−11 (1998) 2278−2324.

35）Breiman, L.: Random forests. Machine Learning 45−1 (2001) 5−32.

36）Hinton, Geoffrey E., Simon Osindero, and Yee-Whye Teh.: A fast learning algorithm for deep belief nets. Neural computation 18−7 (2006) 1527−1554.

37）Hinton, Geoffrey E., and Ruslan R. Salakhutdinov.: Reducing the dimensionality of data with neural networks. science 313−5786 (2006) 504−507.

38）Krizhevsky, A., Sutskever, I., and Hinton, G. E.: ImageNet classification with deep convolutional neural networks. Conf (NIPS, 2012) 1106−1114

39）Bengio, Y., Lamblin, P., Popovici, D., Larochelle, H.: Greedy layer-wise training of deep networks. Conf (Advances in Neural Information Processing Systems, 2007). pp. 153−160.

40）Hinton, G. E., Osindero, S., Teh, Y.-W.: A fast learning algorithm for deep belief nets. Neural Comput. 18 (2006) 1527−1554

41）Krizhevsky, A., Sutskever, I., and Hinton, G. E: ImageNet classification with deep convolutional neural networks. Conf (Advances in Neural Information Processing Systems, 2012). pp.1106−1114.

42）Simonyan, K., Zisserman, A.: Very deep convolutional networks for large-scale image recognition. arXiv:1409−1556 (2014)

43）Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V., Rabinovich, A.: Going deeper with convolutions. arXiv:1409.4842

44）Lin, M., Chen, Q., Yan, S., Network in network. arXiv:1312−4400 (2013)

45）He, K., Zhang, X., Ren, S., Sun, J.: Deep residual learning for image recognition. arXiv:1512.03385 (2015)

46）Gulshan, V., Peng, L., Coram, M., Stumpe, M. C., Wu, D., Narayanaswamy, A., Venugopalan, S., Widner, K., Madams, T., Cuadros, J., Kim, R., Raman, R., Nelson, P. C., Mega, J. L., Webster, D. R.: Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. JAMA 316 (2016) 2402−2410

47）Esteva, A., Kuprel, B., Novoa, R. A., Ko, J., Swetter, S. M., Blau, H. M., Thrun, S., Dermatologist-level classification of skin cancer with deep neural networks. Nature 542 (2017) 115−118

48）Bengio, Y., Simard, P., Frasconi, P.: Learning long-term dependencies with gradient descent is difficult. IEEE Trans Neural Network 5 (1994) 157−166

49）Hochreiter, S., Schmidhuber, J.: Long short-term memory. Neural Computation 9−8 (1997) 1735−1780

50）Cho, K., Van Merrienboer, B., Gulcehre, C., Bahdanau, D., Bougares, ¨ F., Schwenk, H., Bengio, Y.: Learning phrase representations using rnn encoder-decoder for statistical machine translation. arXiv:1406.1078.

(2014)
51) Stollenga, M. F., Byeon, W., Liwicki, M., Schmidhuber, J.: Parallel multi-dimensional LSTM, with application to fast biomedical volumetric image segmentation. Conf (Advances in Neural Information Processing Systems, 2015) pp2998-3006
52) Jia, Y., Shelhamer, E., Donahue, J., Karayev, S., Long, J., Girshick, R., Guadarrama, S., Darrell, T.: Caffe: Convolutional architecture for fast feature embedding. Conf. (Proceedings of the 22nd ACM International Conference on Multimedia, 2014) 675-678.
53) Seide F, Agarwal A. Cntk: Microsoft's open-source deep-learning toolkit, Conf (Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. ACM, 2016) 2135-2135.
54) Abadi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z., Citro, C., Corrado, G. S., Davis, A., Dean, J., Devin, M., Ghemawat, S., Goodfellow, I., Harp, A., Irving, G., Isard, M., Jia, Y., Jozefowicz, R., Kaiser, L., Kudlur, M., Levenberg, J., Mane, D., Monga, R., Moore, S., Murray, D., Olah, C., Schuster, M., Shlens, J., Steiner, B., Sutskever, I., Talwar, K., Tucker, P., Vanhoucke, V., Vasudevan, V., Viegas, F., Vinyals, O., Warden, P., Wattenberg, M., Wicke, M., Yu, Y., Zheng, X.: Tensorflow: Large-scale machine learning on heterogeneous distributed systems. arXiv:1603.04467 (2016)
55) Collobert, R., Kavukcuoglu, K., Farabet, C.: Torch7: A matlablike environment for machine learning. Conf (Advances in Neural Information Processing Systems, 2011)
56) Chen, Tianqi, et al.: Mxnet: A flexible and efficient machine learning library for heterogeneous distributed systems. arXiv preprint arXiv:1512.01274 (2015).
57) Yashiro, S., J. Takatsubo, H. Miyauchi, and N. Toyama: A novel technique for visualizing ultrasonic waves in general solid media by pulsed laser scan. NDT & E International 41-2(2008)137-144.
58) Drai, R., Khelil, M., Benchaala, A.: Time frequency and wavelet transform applied to selected problems in ultrasonics NDE. NDT & E International, 35 (2002) 567-572
59) Mirapeix, J., et al.: Real-time arc-welding defect detection and classification with principal component analysis and artificial neural networks. NDT & e International 40-4 (2007) 315-323.
60) Lawson S W, Parker G A.: Automatic detection of defects in industrial ultrasound images using a neural network Vision Systems: Applications. International Society for Optics and Photonics, 2786 (1996) 37-48.
61) Robini M C, Magnin I E, Benoit-Cattin H, et al.: Two-dimensional ultrasonic flaw detection based on the wavelet packet transform. IEEE transactions on ultrasonics, ferroelectrics, and frequency control, 44-6 (1997) 1382-1394.
62) Zahran O, Al-Nuaimy W.: UTILISING PHASE RELATIONSHIPS FOR AUTOMATIC WELD FLAW CATEGORISATION IN TIME-OF-FLIGHT DIFFRACTION IMAGESC) Conf (ICF11, Italy, 2005)
63) Shitole C S N, Zahran O, Al-Nuaimy W: Combining fuzzy logic and neural networks in classification of weld defects using ultrasonic time-of-flight diffraction. Insight-Non-Destructive Testing and Condition Monitoring, 49-2 (2007) 79-82.
64) Wei, Liang, Zuo-ying Huang, and Pei-wen Que.: Sparse deconvolution method for improving the time-resolution of ultrasonic NDE signals. Ndt & E International 42-5 (2009) 430-434.
65) Sambath, S., P. Nagaraj, and N. Selvakumar.: Automatic defect classification in ultrasonic NDT using artificial intelligence. Journal of nondestructive evaluation 30-1 (2011) 20-28.
66) Zhang, Guang-Ming, Cheng-Zhong Zhang, and David M. Harvey.: Sparse signal representation and its applications in ultrasonic NDE. Ultrasonics 52-3 (2012) 351-363.
67) Guarneri, Giovanni Alfredo, Flávio Neves Junior, and Lúcia Valéria Ramos de Arruda.: Comparative evaluation of artificial neural networks models to classify weld flaws using pulse-echo Ultrasonic signals, Conf (22nd International Congress of Mechanical Engineering (COBEM 2013), Brazil, 2013)
68) Virupakshappa, Kushal, and Erdal Oruklu.: Ultrasonic flaw detection using Support Vector Machine classification. Conf (Ultrasonics Symposium (IUS), 2015 IEEE International.), 2015.
69) Jin, Haoran, et al.: Sparse deconvolution method for ultrasound images based on automatic estimation of reference signals. Ultrasonics 67 (2016) 1-8.
70) Virupakshappa K, Oruklu E.: Ultrasonic flaw detection using Hidden Markov Model with wavelet features

Conf(Ultrasonics Symposium (IUS), 2016 IEEE International. IEEE, 2016) 1-4.
71) Tian, Ye, et al.: A Statistical Framework for Improved Automatic Flaw Detection in Nondestructive Evaluation Images. Technometrics 59-2 (2017) 247-261.
72) Meng M, Chua Y J, Wouterson E, et al.: Ultrasonic signal classification and imaging system for composite materials via deep convolutional neural networks, Neurocomputing, 257 (2017) 128-135.
73) Virupakshappa, Kushal, and Erdal Oruklu.: Performance analysis of deep learning architectures for ultrasonic NDE applications. Conf(Ultrasonics Symposium (IUS), 2017 IEEE International. IEEE, 2017)
74) Wang B, Saniie J.: Ultrasonic flaw detection based on temporal and spectral signals applied to neural network, (Ultrasonics Symposium (IUS), 2017 IEEE International. IEEE, 2017) 1-4.
75) Cruz, F. C., et al.: Efficient feature selection for neural network-based detection of flaws in steel welded joints using ultrasound testing. Ultrasonics 73 (2017) 1-8.
76) Tiwari, Kumar Anubhav, Renaldas Raisutis, and Vykintas Samaitis.: Hybrid Signal Processing Technique to Improve the Defect Estimation in Ultrasonic Non-Destructive Testing of Composite structures. Sensors 17.12 (2017): 2858.
77) K. Supreet Alguri, Chen Ciang Chia and Joel B. Harley,: Model-driven, Wavefield Baseline Subtraction for Damage Visualization using Dictionary Learning, Conf(International Workshop for Structural Health Monitoring (IWSHM) 03/20/2017)
78) Melville, J., Alguri, K. S., Deemer, C., & Harley, J. B.: Structural Damage Detection Using Deep Learning of Ultrasonic Guided Waves. Journal of Engineering mechanics 136-8 (2010) 937-944.
79) Perazzi, Federico, Anna Khoreva, Rodrigo Benenson, Bernt Schiele, and Alexander Sorkine-Hornung.: Learning video object segmentation from static images. Conf (Computer Vision and Pattern Recognition. 2017)
80) Jiang, Y.G., Wu, Z., Wang, J., Xue, X. and Chang, S.F.: Exploiting feature and class relationships in video categorization with regularized deep neural networks. IEEE transactions on pattern analysis and machine intelligence, 40-2 (2018) 352-364.
81) Takahashi, Naoya, Michael Gygli, and Luc Van Gool: Aenet: Learning deep audio features for video analysis. IEEE Transactions on Multimedia 20-3 (2018) 513-524.
82) Sansalone M. Impact-echo : the complete story. ACI Struct J. 94 (1997) 777-86
83) Sansalone, M.J.; William, B.S.: Impact-echo. Nondestructive Evaluation of Concrete and Masonry, (Bullbrier Press: Jersey Shore, PA, USA, 1997)
84) Li, B.; Ushiroda, K.; Yang, L.; Song, Q.; Xiao, J.: Wall-climbing robot for non-destructive evaluation using impact-echo and metric learning SVM. Int. J. Intell. Robot. Appl.1 (2017) 255-270.
85) Igual, J.; Salazar, A.; Safont, G.; Vergara, L. Semi-supervised Bayesian classification of materials with impact-echo signals. Sensors, 15 (2015) 11528-11550
86)Yeh, P.L.; Liu, P.L.: Application of the wavelet transform and the enhanced Fourier spectrum in the impact echo test. NDT Int., 41 (2008) 382–394
87) Igual, J.; Salazar, A.; Safont, G.; Vergara, L. Semi-supervised Bayesian classification of materials with impact-echo signals. Sensors, 15 (2015) 11528-11550.
88) Ye, J.; Takumi, K.; Iwata, M.; Murakawa, M.; Higuchi, T.; Kubota, T.; Yui, T. Noise Reduction Methods for Hammering Impact Acoustic Inspection: An Experimental Comparison. Conf (International Conference on Structural Health Monitoring, Stanford, CA, USA, 1-3 September 2015).
89) Ye, J.; Takumi, K.; Iwata, M.; Murakawa, M. Noise robust hammering echo analysis for concrete structure assessment under mismatch conditions: A sparse coding approach. Conf (proceedings of the IEEE Sensors Applications Symposium (SAS), Glassboro, NJ, USA, 13-15 March 2017) 1-6.
90) Gepperth, A.; Hammer, B.: Incremental learning algorithms and applications. Conf(European Symposium on Artificial Neural Networks (ESANN), Bruges, Belgium, 27-29 April 2016) 357-368
91) Hoi, S.C.H.; Wang, J.; Zhao, P. Libol: A library for online learning algorithms. J. Mach. Learn. Res. 15 (2014) 495-499
92) Peter Stone, Rodney Brooks, Erik Brynjolfsson, Ryan Calo, Oren Etzioni, Greg Hager, Julia Hirschberg, Shivaram Kalyanakrishnan, Ece Kamar, Sarit Kraus, Kevin Leyton-Brown, David Parkes, William Press, AnnaLee Saxenian, Julie Shah, Milind Tambe, and Astro Teller.: Artificial Intelligence and Life in 2030: One Hundred Year Study on Artificial Intelligence: Report of the

2015–2016 Study Panel, (Stanford University, Stanford, CA, September 2016)
93) Chapelle, Olivier, Bernhard Scholkopf, and Alexander Zien.: Semi-supervised learning. *IEEE Transactions on Neural Networks* 20–3 (2009) 542–542.
94) Mann, Gideon S., and Andrew McCallum.: Simple, robust, scalable semi-supervised learning via expectation regularization. Conf (In *Proceedings of the 24th international conference on Machine learning*, ACM, 2007) 593–600
95) Mahdhaoui, Ammar, and Mohamed Chetouani.: Supervised and semi-supervised infant-directed speech classification for parent-infant interaction analysis. *Speech Communication* 53–9 (2011): 1149–1161.
96) Papandreou, George, Liang-Chieh Chen, Kevin P. Murphy, and Alan L. Yuille.: Weakly-and semi-supervised learning of a deep convolutional network for semantic image segmentation. Conf (In *Proceedings of IEEE international conference on computer vision*. 2015) 1742–1750
97) Kingma, Diederik P., and Max Welling. Stochastic gradient VB and the variational auto-encoder. Conf (*Second International Conference on Learning Representations, ICLR*. 2014)
98) Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., ... & Bengio, Y.. Generative adversarial nets. Conf (*Advances in neural information processing systems*, 2014) 2672–2680
99) Souly, Nasim, Concetto Spampinato, and Mubarak Shah: Semi Supervised Semantic Segmentation Using Generative Adversarial Network. Conf (In *Proceedings of IEEE International Conference on Computer Vision (ICCV)*, 2017) 5689–5697.
100) Yoon, Andre S., Taehoon Lee, Yongsub Lim, Deokwoo Jung, Philgyun Kang, Dongwon Kim, Keuntae Park, and Yongjin Choi.: Semi-supervised Learning with Deep Generative Models for Asset Failure Prediction. *arXiv preprint arXiv:1709.00845* (2017).