IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received November 28, 2019, accepted December 7, 2019, date of publication December 13, 2019,
date of current version December 26, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2958962

A Survey on Deep Learning Empowered IoT Applications

XIAOQIANG MA[1], (Member, IEEE), TAI YAO[1], MENGLAN HU[1], YAN DONG[1], WEI LIU[1], FANGXIN WANG[2], AND JIANGCHUAN LIU[2], (Fellow, IEEE)

[1]School of Electronic Information and Communications, Huazhong University of Science and Technology, Wuhan 430074, China
[2]School of Computing Science, Simon Fraser University, Burnaby, BC V5A 1S6, Canada

Corresponding author: Fangxin Wang (fangxinw@sfu.ca)

This work was supported in part by the National Natural Science Foundation of China under Grant 91538203, Grant 61977064, Grant 61702204, Grant 61872416, Grant 61671216, Grant 61871436, Grant 61872415, and Grant 61602214, and in part by the Fundamental Research Funds for the Central Universities of China under Grant 2019kfyXJJS017.

ABSTRACT The Internet of Things (IoT) is widely regarded as a key component of the Internet of the future and thereby has drawn significant interests in recent years. IoT consists of billions of intelligent and communicating ''things'', which further extend borders of the world with physical and virtual entities. Such ubiquitous smart things produce massive data every day, posing urgent demands on quick data analysis on various smart mobile devices. Fortunately, the recent breakthroughs in deep learning have enabled us to address the problem in an elegant way. Deep models can be exported to process massive sensor data and learn underlying features quickly and efficiently for various IoT applications on smart devices. In this article, we survey the literature on leveraging deep learning to various IoT applications. We aim to give insights on how deep learning tools can be applied from diverse perspectives to empower IoT applications in four representative domains, including smart healthcare, smart home, smart transportation, and smart industry. A main thrust is to seamlessly merge the two disciplines of deep learning and IoT, resulting in a wide-range of new designs in IoT applications, such as health monitoring, disease analysis, indoor localization, intelligent control, home robotics, traffic prediction, traffic monitoring, autonomous driving, and manufacture inspection. We also discuss a set of issues, challenges, and future research directions that leverage deep learning to empower IoT applications, which may motivate and inspire further developments in this promising field.

INDEX TERMS Internet of Things, deep learning, smart healthcare, smart home, smart transportation.

I. INTRODUCTION

The rise of Internet-of-Things (IoT) technology has brought prosperity to a myriad of emerging applications on various mobile and wireless platforms including smart phones [1], sensor networks [2], unmanned aerial vehicles (UAV) [3], [4], cognitive smart systems [5], and so on. To develop effective IoT applications, we may typically follow a work-flow model, which consists of five components: question formulation, data collection, data analysis, visualization, and evaluation [6]. Among them, data analysis is a critical and computational intensive part wherein traditional technologies generally combine professional knowledge with machine learning (e.g., logistic regression, support vector machine, and random forest) to figure out classification or regression problems (e.g. the traffic condition prediction with support vector machine (SVM) [7], car tracking with Kalman filter and ridge regression [8], delivery time estimation with Gaussian mixture model (GMM) [9], and localization with SVM [10]). However, as the human society steps into the ''Big Data'' era, such conventional approaches are not sufficiently powerful to process the massive, explosive, and irregular data collected from ubiquitous and heterogeneous IoT data sources. Almost all traditional systems rely on specially designed features, and the performance heavily depends on the prior knowledge of specific fields. Most of learning techniques applied in such systems normally utilize shallow architectures, which have very limited modeling and representational power. As such, a more powerful analytical tool is highly desirable to fully unleash the potentials of the invaluable raw data generated in various IoT applications.

The associate editor coordinating the review of this manuscript and approving it for publication was Chin-Feng Lai.

The recent breakthroughs in deep learning and hardware design have enabled researchers to train much more powerful models, which highly empower many applications such as crowdsourced delivery [11], network caching [12], energy management [13] and edge computing [14]. In the following we highlight the advantages of deep learning as compared with traditional machine learning methods, which demonstrates the benefits to apply deep learning in IoT applications.

- Deep learning incorporates deeper neural network architectures, which is able to extract more complex hidden features (such as temporal and/or spatial dependencies) and characterize more intricate problems. Different from those traditional simple learning methodologies, deep learning has more powerful capabilities in generalizing the complicated relationship of massive raw data in various IoT applications.
- Deep learning is able to fully take advantages of the massive yet invaluable data resource. The data processing ability typically depends on the depth and the particular architectures of learning models, such as convolutional architectures; hence, deep learning based models can mostly perform better on large scale data, while simple learning models may be easily over-fitted when dealing with the deluge of data.
- Deep learning is a kind of end-to-end learning method that is able to automatically learn how to directly extract effective features from the raw data without the involvement of the time-consuming and laborious hand-crafted feature specification.

While a lot of efforts have been made in the past few years, the whole area of leveraging deep learning in IoT applications is still at an infant stage. A few articles which survey the applications of deep learning in IoT domains have been presented in the literature. Alsheikh *et al.* [15] mainly reviewed papers in applying machine learning in wireless sensor networks (WSN). In [16], the authors focused on the survey of applying deep learning techniques for healthcare applications. Another work [17] surveys state-of-the-art deep learning methods and their applicability in the IoT applications, with an emphasis on big data and streaming data analytics. The authors in [18] present a comprehensive survey of commonly used deep learning algorithms and discuss their applications towards making manufacturing smart. Nevertheless, all these existing survey articles only focus on relatively partial IoT fields. A survey that comprehensively reviews deep learning for a variety of IoT applications is still absent. Therefore, we believe that it is the right time to review the existing literature and to motivate future research directions.

To this end, this article summarizes the up-to-date research progresses and trends on leveraging deep learning tools to empower IoT applications. We put emphasis on four representative IoT application scenarios, including smart healthcare, smart home, smart transportation, and smart industry. We aim to reveal how deep leaning can be applied to enhance IoT applications from diverse perspectives. A main thrust on this topic is to seamlessly merge the two disciplines of deep learning and IoT, resulting in a broad spectrum of novel designs in IoT applications, such as health monitoring, disease analysis, indoor localization, intelligent control, home robotics, traffic prediction, traffic monitoring, autonomous driving, manufacture inspection and fault assessment. We also discuss the issues, challenges and future research directions for applying deep learning in IoT applications. All these insights may motivate and inspire further developments in this promising field.

The rest of the paper is organized as follows: Section II introduces classic deep learning models employed in the following sections, including Restricted Boltzmann Machines (RBMs), Autoencoder, Convolutional Neural network (CNN), and Recurrent Neural Network (RNN). Section III surveys the latest deep learning based IoT applications in four major application scenarios. Section IV outlines challenges and opportunities for leveraging deep learning in IoT applications. Section V concludes the article.

II. OVERVIEW OF DEEP LEARNING METHODS

Deep learning requires powerful computation resources and is known to be time-consuming during the model training phase, which is one of the major challenges in the before. The development of specialized hardware (e.g., powerful GPUs) and efficient training algorithm makes it possible to analyze complex problems and process data by building deep models rather than traditional methods. Hence, deep learning has now become a promising data processing and modeling approach in the era of big data.

Similar to traditional machine learning, deep learning can be divided into two scenarios: unsupervised learning (models trained with unlabeled data) and supervised learning (models trained with labeled data). In the following, we will briefly introduce four deep learning models, including two unsupervised learning models, restricted Boltzmann machines (RBMs) and Autoencoder, as well as two supervised learning models, Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs).

A. UNSUPERVISED LEARNING

The collection of unlabeled data can be carried out easily. In order to deal with massive unlabeled data, unsupervised learning must be applied as a supplement of conventional learning methods. The training can be conducted with stacked RBMs or stacked autoencoders for stable initialization, back propagation, and global fine-tuning.

1) RESTRICTED BOLTZMANN MACHINES

Restricted Boltzmann machines (RBMs) [19] are probabilistic graphical models that can be interpreted as stochastic neural networks. RBMs consist of m visible units to represent observable data and n hidden units to capture collections between observed variables, providing us a stochastic representation of the output. Fig. 1 shows a two-level RBM with m visible variables and n hidden variables. RBMs are successful

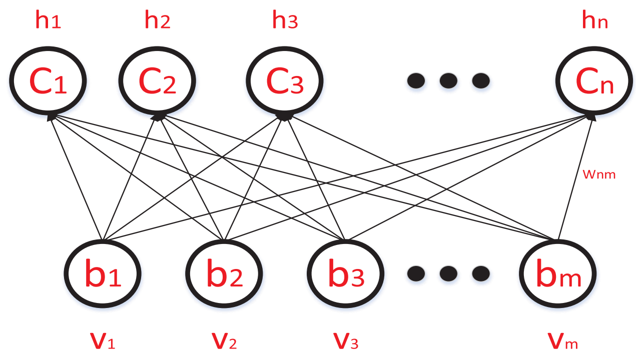


FIGURE 1. An RBM with *m* visible and *n* hidden variables.

in dimensionality reduction and collaborative filtering [20]. A Deep Belief Network (DBN) forms a deep learning model by stacking RBMs [21], which is trained in a layer-by-layer manner using a greedy learning algorithm, and the contrastive divergence (CD) method is applied to update the weights. Neural networks are prone to trap in the local optima of a non-convex function, resulting in poor performance [22]. DBN incorporates both unsupervised pre-training and supervised fine-tuning methods to construct the models: the former intends to learn data distributions with unlabeled data and the latter aims to obtain an optimal solution through fine tuning with labeled data [23].

2) AUTOENCODER

An autoencoder [24] is a neural network trained to copy its input to its output. Compared to RBMs, an autoencoder consists of three layers including an input layer, a hidden layer, and an output layer. The hidden layer describes a code used to represent the input, and its output is a reconstruction of the input. Basically, the network consists of two major components: an encoder function **f** which extracts the dependencies of the input, and a decoder **g** function which produces a reconstruction. Autoencoder is trained by minimizing the error between the input and output. Fig. 2 shows a brief architecture of an autoencoder and a concrete example. Like RBMs, a deep model can be constituted through a stack of autoencoders in a layer-by-layer manner. The hidden layer of a well-trained autoencoder is fed as the input layer of another autoencoder, and iteratively a multi-layers model is formed. The variants of autoencoder include sparse autoencoder [25], denoising autoencoder [26], and contractive autoencoder.

B. SUPERVISED LEARNING

Unlike unsupervised learning, a labeled training set is a part to build the system model in supervised learning. The model learns the relationship between the input, output, and system parameters. The major method used in supervised learning is the back propagation algorithm [22].

1) CONVOLUTIONAL NEURAL NETWORKS (CNNs)

CNN is a specialized kind of neural network for processing data that has a known, grid-like topology [24]. CNNs were

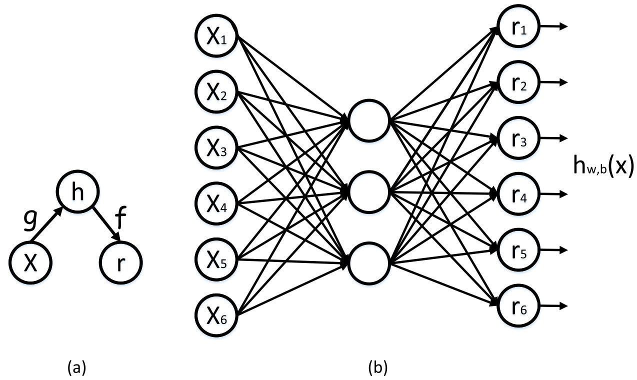


FIGURE 2. (a) The general structure of an autoencoder, and (b) a specific autoencoder structure diagram which has 6 input data.

firstly inspired by a concept called Receptive Field which comes from the study of cat's visual cortex [27]. Convolution leverages three important ideas that can help improve a machine learning system: sparse interactions, parameter sharing, and equivariant representations. The basic CNN architecture is made up by one convolutional and pooling layer, optionally followed by a fully connected layer for classification or prediction. In contrast to traditional neural networks, CNN efficiently decreases the number of parameters in nets and the effect of gradient diffusion problem, which means that we can successfully train a deep model containing more than 10 layers using CNNs.

For example, AlexNet [28] contains 9 layers, VGGNet [29] contains 11-19 layers, InceptionNet [30] from Google contains more than 22 layers, and ResNet [31] from Microsoft even contains 152 layers. Fig. 3 shows a general architecture of traditional CNNs called LeNet [32].

2) RECURRENT NEURAL NETWORKS (RNNs)

RNN is a family of neural networks for processing sequential data. RNNs are practical to scale much longer sequences than networks without sequence-based specialization. Many recurrent neural networks apply equation or a similar equation $h^{(t)} = f(h^{t-1}, x^{(t)}; \theta)$ to define the values of their hidden units, illustrated in Fig. 4 [24]. From the network structure, we can observe that RNNs can remember the previous information and utilize it to influence the output of the subsequent nodes. However, RNNs are restricted by looking back only a few steps, due to the gradient diffusion problem and long-term dependencies. To solve these problems, new approaches like LSTM (Long Short-Term Memory) [33] and GRU (Gated Recurrent Unit) [34] have been proposed, modeling the hidden state to decide what to keep in the previous and current memory. These variants can efficiently capture the long-term dependencies and lead to a stronger capacity to understand the language. Different from CNN which processes spatially continuous data, RNN focuses on the connections between temporally continuous data. Therefore, RNN is mostly employed in the natural language processing (NLP) field [35]–[37].

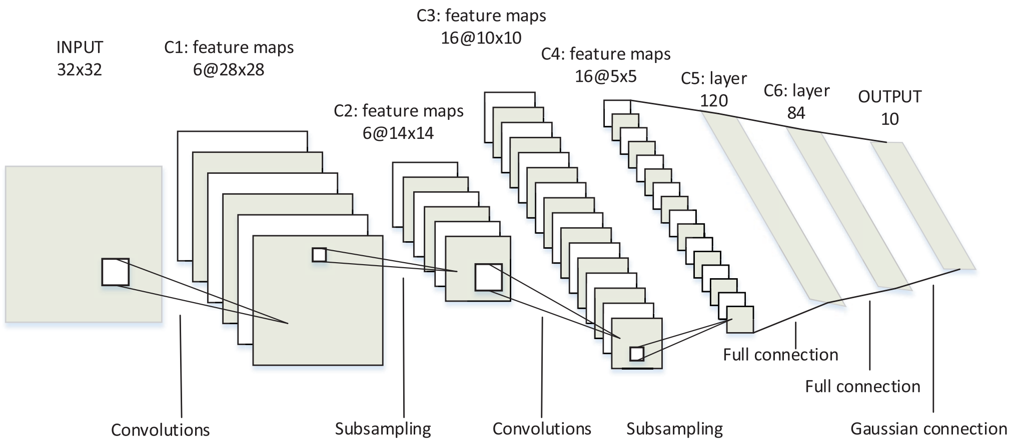


FIGURE 3. The architecture of LeNet 5. Each plane represents a feature map. The little white boxes are the keys for convolutional neural networks, which are called kernels. From the figure we can see that, contrast to full connection layers, the convolutional layers focus more on local relationships.

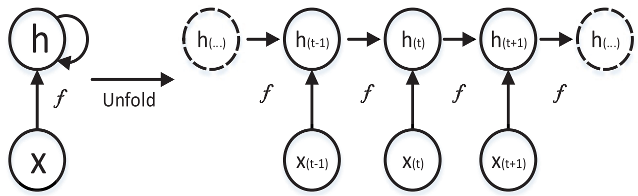


FIGURE 4. A general structure of unfolding recurrent neural networks without the output.

III. WHEN IoT MEETS DEEP LEARNING

Thanks to recent breakthroughs in hardware and theories, the application of deep learning in IoT are rapidly emerging and developing. In this section, we provide a comprehensive overview of the latest development in four representative IoT areas. Our discussion starts from healthcare applications, which include health monitoring and disease analysis. We next discuss how deep learning can be employed to process enriching sensing signals and make our home smarter. Third, we introduce various deep learning applications in smart transportation, mainly in three fields: traffic prediction, traffic monitoring and autonomous driving. Finally, smart industry is another promising field where we focus on manufacturing inspection and related applications.

A. SMART HEALTHCARE

Gaining knowledge and actionable insights from complex, high-dimensional, and heterogeneous biomedical data remains a key challenge in transforming health care. Nowadays, IoT combined with deep learning has been employed in providing healthcare and well-being solutions for individuals and communities, and deep learning approaches could be the vehicle for translating big biomedical data into improved human health [38]. We summarize some typical deep learning-based smart healthcare applications in Table 1.

1) HEALTH MONITORING

Nowadays, sensor-equipped smartphones and wearables customarily enable a variety of mobile APPs for health monitoring [49], [50]. To implement such applications, people utilize Human Activity Recognition (HAR) to identify human activities and analyze health conditions [16]. However, the underlying representative features hidden in the massive raw data calls for more effective extraction model for identification. Applying the advance of deep learning in activity recognition opens a promising opportunity towards this problem. Hammerla et al. [39] build CNNs and LSTM to analyze the movement data respectively and combine the results to make a better prediction of freezing gaits in Parkinson disease patients. Zhu et al. [41] apply the data from triaxial accelerometers and heart rate sensors to obtain promising results in predicting Energy Expenditure (EE) with a CNN model, which helps to relieve chronic diseases. Hannun et al. [42] train a 34-layer convolutional neural network which maps a sequence of ECG samples to a sequence of rhythm classes. The performance exceeds that of board certified cardiologists in detecting a wide range of heart arrhythmias from electrocardiograms recorded with a single-lead wearable monitor. Gao et al. [40] propose a novel deep learning architecture recurrent 3D convolutional neural network (R3D). R3D extracts effective and discriminative spatial-temporal features for action recognition, which enables the capturing of long-range temporal information by aggregating the 3D convolutional network entries to serve as an input to the LSTM architecture.

With the prevalence of wearable devices, we can monitor our health state and standardize our way of life at any time. It is a challenge to directly deploy deep learning modules on low-power wearable devices due to their limited resources. Ravì *et al.* [51] utilize spectral domain preprocessing before the data are passed onto the deep learning framework so as to optimize real-time on-node computation in resource-limited devices. Tang *et al.* [52] explore two ways and successfully integrate deep learning with low-power IoT products.

2) DISEASE ANALYSIS

Medical image classification and analysis is an important topic in healthcare. Following the success in computer vision,

TABLE 1. Summary of deep learning-based smart healthcare applications.

Category	Application	Model	Remark
Health Monitoring	Human activity recognition (HAR) [39]	CNN + LSTM	Compare different models using three representative datasets that contain movement data captured with wearable sensors
	Human action monitoring [40]	3D CNN + LSTM	A deep learning architecture named as recurrent 3D convolutional neural network (R3D) to extract effective and discriminative spatial-temporal features to be used for action recognition
	Energy expenditure (EE) estimation [41]	CNN	Use CNNs to automatically detect important features from data collected from triaxial accelerometer and heart rate sensors
	Arrhythmia detection and classification [42]	CNN	Develop a 34-layer deep neural network to classify 10 arrhythmias as well as sinus rhythm and noise from a single-lead ECG signal, with high diagnostic performance similar to that of cardiologists
Disease Analysis	Knee cartilage segmentation [43]	CNN	Infer a hierarchical representation of low-field knee MRI scans that fosters categorization
	Detection of diabetic detinopathy in retinal fundus photographs [44]	CNN	High sensitivity and specificity
	Smart personal health advisor [45]	CNN	A smart personal health advisor (SPHA) for comprehensive and intelligent physiological and psychological health monitoring and guidance.
Other	Pill image recognition [46]	CNN	A mobile vision system for recognizing unconstrained pill images
	Skin lesion classification [47]	CNN	A deep CNNs-based model to predict whether a image of a skin lesion is either benign or malignant
	Medicine recognition [48]	Faster R-CNN	ST-Med-Box can assist chronic patients in taking multiple medications correctly and avoiding in taking the wrong medications

deep learning has been widely used in assisting disease image analysis [45], [53]. CNNs are used to infer a hierarchical representation of low-field knee MRI scans to automatically segment cartilage and predict the risk of osteoarthritis [43]. Another work [44] uses CNNs to identify diabetic retinopathy in retinal fundus photographs, obtaining high sensitivity and specificity over about 10, 000 test images with respect to certified ophthalmologist annotations.

In addition to medical image recognition, deep learning has been employed in other applications. For instance, Zeng *et al.* [46] present a deep-learning based pill image recognition model which helps to identify unknown prescription pills using smartphones. Lopez *et al.* [47] propose a deep-learning-based approach to classify a dermotropic image which contains a skin lesion as malignant or benign. A ubiquitous healthcare framework called UbeHealth is proposed to address the challenges in terms of network latency, bandwidth, and reliability [54]. Chang *et al.* [48] propose an intelligent medicine recognition system called ST-Med-Box based on deep learning. ST-Med-Box can assist chronic patients in taking multiple medications correctly and avoiding in taking wrong medications; it also provides other medication-related functionalities such as reminders to take medications on time, medication information, and chronic patient information management.

B. SMART HOME

Smart home enables the innerconection of ubiquitous smart home devices and is the technical convergence and services through home networking to achieve a better quality of living [55]. In the recent years, a myriad of systems has been developed to apply deep learning techniques in various smart home applications, as summarized in Table 2.

1) INDOOR LOCALIZATION

With the proliferation of mobile devices, indoor localization gradually becomes a critical research issue since it is not viable to employ Global Positioning System (GPS) in indoor environments. Indoor localization enables numerous services in smart home, such as wireless intruder detection, elder monitoring, and baby monitoring, yet it faces a lot of propagation challenges like multi-path effect, fading, and delay distortion. High accuracy and short processing time are indispensable performance indicators while designing an indoor localization system. Fingerprinting-based indoor localization is an effective method to satisfy the above requirements. RSSI (Received Signal Strength Indication) based fingerprints are known to be unstable and inaccurate, and the more powerful Wi-Fi Channel State Information (CSI) have become the most widely adopted fingerprints in current systems. In addition, traditional positioning systems are based on such methods as K nearest neighbors (KNN) [62], Bayesian model [63], SVM [64], and compressive sensing [65], which are not suitable for dealing with massive data. To this end, people began to resort to deep neural networks.

Gu *et al.* [56] propose a novel algorithm called Semi-supervised Deep Extreme Learning Machine (SDELM), which takes the advantages of semi-supervised learning, deep learning, and extreme learning machine. This approach achieves satisfactory performance on the localization and reduces the calibration effort with the full use of unlabeled data. Mohammadi *et al.* [57] propose a semisupervised deep reinforcement learning (DRL) model based on Bluetooth low energy signal strengths. This model utilizes variational autoencoders as the inference engine for generalizing optimal policies. Wang *et al.* [58] utilize 4-layer RBMs to process the raw CSI data to obtain the locations. Yet the proposed

TABLE 2. Summary of deep learning-based smart home applications.

Category	Application	Model	Remark
Indoor Localization	Semi-supervised Wi-Fi based localization [56]	DNN	Combines semi-supervised learning, deep learning, and extreme learning machine (ELM)
	Bluetooth low energy signal strength based indoor localization [57]	DRL	A semi-supervised deep reinforcement learning model that leverages both labeled and unlabeled data
	CSI-based fingerprinting for indoor localization [58]	RBM	Device-oriented, one AP
	Device-free wireless localization and activity recognition [59]	RBM	Device-free, nine APs
Home Robotics	Autonomous navigation [60]	CNN	Use CNN to identify markers or objects from images and videos
	Hand-eye coordination for robotic grasping [61]	CNN	The learning algorithm is able discover unconventional and non-obvious grasping strategies

system considers a device-oriented approach, which would not work if people have no cell phones or they refuse to connect their phones with APs. To this end, Wang *et al.* [59] develop a device-free approach based on an observation that APs receive different data when people stand at different locations. They design a 4-layer RBM model to extract features from the raw CSI data and select random forests (RF) to classify the locations by these features. In addition, they employ a contaminant estimation step to eliminate the error of CSI values in a fixed place due to multi-path effect caused by opening a window or door. Nine APs are employed to collect data related to people's locations, and a wavelet filter is utilized to preprocess the raw data. With the multi-faceted interaction, results are more robust. The system can even recognize people's activities like bow and walk, or gestures like hand-clap and wave hand.

2) INTELLIGENT CONTROL

Nowadays, home appliances can connect to the Internet and provide intelligent services. Li and Lin [66] utilize WSNs and power line communications (PLCs) to implement a smart home control network. To reduce the impact of wireless interference on the control network and the unnecessary energy consumption, an isolated WSN with one coordinator, which is integrated into the PLC transceiver, is established in each room. The coordinator is responsible for transferring environmental parameters obtained by WSNs to the management station via PLCs. The control messages for home appliances are directly transferred by PLCs rather than WSNs. The user interface is also an important research field for better user experience. The authors in [67] propose a gesture-based user interface for the development of a smart home system.

Nowadays, deep learning techniques have shown great success in digital personal assistant products such as Microsoft's Cortana, Apple's Siri, Amazon Alexa, and Google Assistant [68]. Such dialogue system based products would function as the next-generation smart home controller.

3) HOME ROBOTICS

With sensors, actuators and databases equipped, home robots can perform various tasks in home environments. In general, home service robots should have the key functionalities including the localization, navigation, map building, human-robot interaction, object recognition, and object handling [69]. Robotic navigation in GPS-denied environments requires case-specific approaches for controlling a mobile robot to any desired destinations. In [60], a new approach for autonomous navigation to identify markers or objects from images and videos is presented, using pattern recognition and machine learning techniques such as CNNs. Computational intelligence techniques are implemented along with the robot operating system and object positioning to navigate towards these objects and markers by using the RGB-depth camera. Multiple potential matching objects detected by the robot with deep neural network object detectors are displayed on a screen installed on the assisted robot to improve and evaluate Human-Robot Interaction (HRI). To improve the hand-eye coordination for the object handling, Levine *et al.* [61] train a large convolutional neural network to predict the probability that the task-space motion of the gripper results in successful grasps only using monocular camera images, independently of the camera calibration or the current robot pose.

C. SMART TRANSPORTATION

In this section, we mainly present how deep learning has been applied in various smart transportation applications, which is summarized in Table 3.

1) TRAFFIC FLOW PREDICTION

Traffic flow prediction is a fundamental problem in transportation modeling and management as well as intelligent transportation system design, which nowadays heavily depends on the historical and real-time traffic data collected from all kinds of sensors, including inductive loops, cameras, crowd sourcing, social media and so on. To efficiently utilize such massive heterogeneous data, classical machine learning methods, e.g., SVM, would consume a lot of time and power-consuming computation resources. In addition, hand-engineered features are not enough for a satisfying accuracy due to the limitation of related prior knowledge. The authors in [7] propose an online-SVR method for short traffic flow prediction in typical and atypical conditions, where several SVM models need to be formed and consume a lot of memory resources.

TABLE 3. Summary of deep learning-based smart transportation applications.

Category	Application	Model	Remark
Traffic Prediction	Traffic flow prediction [70]	DBN	A deep architecture that consists of two parts, a DBN at the bottom for unsupervised feature learning, and a multitask regression layer at the top for supervised prediction
	Traffic flow prediction [71]	SAE	A stacked autoencoder (SAE) model is used to learn generic traffic flow features
	Short-term traffic prediction [72]	LSTM	A LSTM network considers temporal-spatial correlation in traffic system via a two-dimensional network
	Crowd flow prediction [73]	CNN	A deep spatio-temporal residual network (ST-ResNet) based on unique properties of spatio-temporal data
Traffic Monitoring	Real-time object detection [74]–[76]	RPN + Faster R-CNN [74], YOLO [75], SSD [76]	Improvement on speed and detection accuracy
	Object tracking [77]	CNN	A novel CNN architecture and a structural loss function that handles multiple input cues, and a tracking-tailored SGD algorithm
	Object tracking [78], [79]	RNN	Track and classify objects in complex and only partially-observable, real-world scenarios from both a static sensor as well as a moving platform
	Road accident detection [80]	SDAE	Learn feature representation from the spatio-temporal volumes of raw pixel intensity
Automonous Driving	Real-time object detection for autonomous driving [81]	CNN	Aims at small, fast, energy efficient, and accurate models
	End-to-end learning for Self-Driving Cars [82]	CNN	Map raw pixels from a single front-facing camera to steering commands
	Learning driving from video datasets [83]	FCN-LSTM	Learn from large-scale crowd-sourced vehicle action data

Recently, deep learning has drawn major attention from both academia and industry due to its ability to extract inherent features from data and exploit the rich amount of traffic data. Huang et al. [70] propose a DBN model to capture enough features from each part of road traffic networks. With the idea of multitask learning, these features from related roads and stations are grouped to explore the nature of the whole road traffic network and predict traffic flow. Lv *et al.* [71] propose an SAE (Stack of Autoencoders) model to extract features from historical data for prediction with these features. A lot of works have focused on utilizing deep learning for traffic and crowd flow prediction [72], [84].

2) TRAFFIC MONITORING

One of the most attractive research fields in smart transportation is the development automated traffic monitoring systems, which play an important role in both reducing the workload of human operators and warning drivers of dangerous situations [89], [90]. Traffic video analytics has become an important part of intelligent traffic monitoring systems. In the following we present how deep learning is applied to traffic video analytics from the three perspectives: object detection, object tracking, and face recognition.

Object detection has been applied in a wide range of scenarios, such as pedestrian detection, on-road vehicle detection, and unattended object detection. Applying the deep convolutional neural network and multi-scale strategy has significantly improved the accuracy and speed [74]–[76]. Ren *et al.* [74] introduce a region proposal network (RPN) that shares full-image convolutional features with detection network, thus enabling nearly cost-free region proposals. Redmon *et al.* [75] recognize frame object detection as a regression problem to spatially separate bounding boxes and associate class probabilities. Liu *et al.* [76] discretize the output space of bounding boxes into a set of default boxes over different aspect ratios and scale per feature map location.

Object tracking is intended to locate a target in a video sequence and give its location in the first frame, which has been applied in surveillance systems. It is important to automatically track suspected people or target vehicles for safety monitoring, urban flow management, and autonomous driving [77], [91]. Vincent *et al.* [91] explore an original strategy for building deep networks based on stacking layers of denoising autoencoders, which are trained locally to denoise corrupted versions of their inputs. Li *et al.* [77] present an efficient and robust tracking algorithm by using a single CNN for learning effective feature representations of target object. To directly map from raw sensor input to object tracks in sensor space without requiring any feature engineering or system identification, the end-to-end object tracking approach has been proposed, where recurrent neural networks (RNN) is used [78], [79]. Singh and Mohan [80] propose a framework for automatic detection of road accidents in surveillance videos, which uses a stacked denoising autoencoder (SDAE) to learn feature representation from the spatio-temporal volumes of raw pixel intensity instead of traditional hand-crafted features. Face recognition and detection techniques [92]–[94] can be used to identify and track drivers and pedestrians.

3) AUTONOMOUS DRIVING

Autonomous driving is a crucial part of city automation. There are two major paradigms for vision-based autonomous

TABLE 4. Summary of deep learning-based smart industry applications.

Category	Application	Model	Remark
Manufacture Inspection	Robust inspection system [85]	CNN	Leverage fog computing to offload the computation burden to the fog nodes to deal with big data in real time
	Surface integration inspection [86]	CNN	Troposed method has several advantages in time and cost saving and shows higher performance than traditional manpower inspection system
Fault Assessment	Fault diagnosis and identification [87]	Wavelet-based CNN	Wavelet transform is used to featureless fault diagnosis
	Transformer fault diagnosis [88]	Continuous Sparse Autoencoder (CSAE)	Gaussian stochastic unit is added into an activation function to extract nonlinear features of the input data

driving systems, the mediated perception approaches and behavior reflex approaches [95]. Systems based on mediated perception approaches compute a high-dimensional world representation. The idea of mediated perception approaches recognizes multiple driving-relevant objects [96], [97], such as lanes, traffic signs, traffic lights, cars, pedestrians, etc. Mediated perception approaches gain the state-of-the-art achievement in autonomous driving. However, most of these systems rely on high precision instruments,and bring unnecessarily high complexity and cost. Currently, autonomous driving systems focus more on real-time inference speed, small model size, and energy efficiency [81]. These self-driving systems are trained by the driving videos to learn a map from input images to driving behaviors or constructs a direct map from the sensory input to a driving action. The authors in [82] train a convolutional neural network to map raw pixels from a single front-facing camera directly to steering commands. Inspired by language models, The authors in [83] put forward a learning-based approach which trains an end-to-end FCN-LSTM network to predict multi-modal discrete and continuous driving behaviors. The system learns from Long-term Recurrent Convolutional Network [98] and extracts the spatial and temporal connections of driving video.

D. SMART INDUSTRY

Industry 4.0, or in other words, the smart industry, represents the latest trend of the manufacturing revolution. In the era of smart industry, explosive data produced in manufacture can be analyzed to empower the computers and manipulators with human-like abilities [85]. In the following part, we mainly focus on how deep learning is implemented in some specific industrial applications, which are summarized in Table 4.

1) MANUFACTURE INSPECTION

In order to accurately inspect and assess the quality of products, various visual inspection approaches, many of which are based on traditional machine learning techniques, have been proposed to extract representative features with expert knowledge so as to detect product defects in large scale production [99]. Recently, deep learning has become a powerful tool for visual inspection. The authors in [85] propose a deep-learning-based classification model to implement a robust inspection system. A CNN-based system is adapted to the fog computing environment, which significantly improves its computing efficiency. A generic CNN-based approach is proposed in [86] to extract patch features and predict defect areas via thresholding and segmenting for the tasks of surface integration inspection.

2) FAULT ASSESSMENT

In order to implement smart manufacturing, it is crucial for a smart factory to monitor machinery conditions, identify incipient defects, diagnose root cause of failures, and then incorporate the information into manufacturing production and control. In [87], a wavelet-based CNN is proposed for automatic machinery fault diagnosis. The wavelet transform is used to transfer one-dimensional vibration signal into a two-dimensional one which is then fed into CNN. In [88], a continuous sparse auto-encoder (CSAE) is presented by adding Gaussian stochastic unit into an activation function to extract nonlinear features of the input data. In [100], a sparse filtering based two-layer neural network model is investigated for unsupervised feature learning, which is used to learn representative features from the mechanical vibration signals.

IV. CHALLENGES AND OPPORTUNITIES

In the above section we have comprehensively surveyed the state-of-arts on applying deep learning in various IoT domains. Nevertheless, the relevant research work is still at an early stage while many key issues need to be addressed in future efforts. In this section, we first identify four critical challenges for implementing deep learning in IoT applications. We then point out two opportunities brought by combining deep learning with IoT technologies.

A. CHALLENGES

As shown in the recent literature, data processing and analysis using deep learning delivers satisfactory performance. Yet a lot of issues have yet to be addressed when leveraging deep learning in IoT applications.

1) DATA COLLECTION

The performance of deep learning methods relies on data sources. Without sufficient clean data, the deep model cannot play a role, even if the architecture of the model is well designed. Therefore, how to implement the data collection equipment is a critical research issue. The number of sensors used and the way to deploy the sensors impact on the quality of data collected. The information contained in the data is actually the key to solve problems. It is necessary to design a data collection module for the whole IoT application work-flow. For example, Li et al. [101] deliberately design a picture collection module *DeepCham* to improve the identification accuracy of the model. Actually, it implements the idea of crowdsourcing in the data collection module. A cost-effective, reliable, and trustworthy data collection paradigm plays an important role in developing practical deep learning based IoT applications.

2) MODEL TRAINING

Training a deep network demands cumbersome tasks. As we know, the depths determine the capacity of a deep learning network to extract key features. However, the gradient vanishment problem appears when models grow deeper, which deteriorates the performance. To this end, Hinton et al. [23] propose an approach to pre-train models by stacking RBMs. In addition, the ReLU function applied as a substitute for the sigmoid function also contributes to the mitigation of the gradient vanishment problem. Overfitting is another serious problem that we face in training deep models. The key solution is to enter more data or reduce parameters of the model. One effective method is using convolutional kernels to reduce the number of parameters, and employing the dropout [102] is also an alternative. Moreover, in recent years, a major breakthrough has been made in convolutional neural networks [28]–[31] and the number of layers of CNN models has been increasing from 5 to more than 200. Methods mentioned in these classical convolutional neural networks (like implementing smaller convolutional kernels or batch normalization) can be valid when we utilize the deep learning algorithm to deal with problems in wireless network filed.

3) HARDWARE LIMITATION

Deep learning is a powerful tool for processing big data, resulting in its high hardware requirements. How to implement a deep model of a resource-limited embedded device is still a challenge. So far, there are two types of research aiming to solve the problem. One is only to treat end devices (like a smart phone) as data collectors. All data are transferred to resourceful servers to be analyzed. However, in this process, we may incur data disclosure, network failure, and other issues. An alternative solution is to reduce the complexity of the networks with slight performance degradation, such that some learning tasks can be performed at end devices.

4) SYSTEM DESIGN

There emerges a trend to design a cloud-edge learning system that spans edge devices and the cloud. A cloud-edge system can leverage the edge to reduce latency, improve the safety and security, and implement intelligent data retention techniques [103]. Furthermore, it can leverage the cloud to share data across edge devices, train sophisticated computation-intensive models, and take high-quality decisions [104]. Recently, there have been some studies on the combination of deep learning and edge computing [105]–[107]. Edge devices can be highly heterogeneous in terms of resource capabilities and software platforms, which make it complicated for application development. The update cycles of the hardware and software to edge devices are much slower than in the cloud. It is inconvenient to store all collected data because of the storage capacity slowing down [104]. It is highly desirable to address these problems and build a robust cloud-edge learning system.

B. OPPORTUNITIES

Despite the challenges, there are still opportunities in applying deep learning to solve IoT problems:

- Deep learning liberates our thought. In the past, we may hesitate to step into some unknown areas and encounter difficulties when we carry out some research due to the limitation of related professional knowledge. Now we can achieve some guess without the worry of data analysis tools. Deep learning gives us the ability to obtain and process data information. It means that we can boldly start more research and may promote the process of science and technology.
- Deep architectures have a strong ability in the representation of learning. Features that describe characteristics of input data used to be designed manually in traditional methods. Deep learning allows machines to design features by themselves. With the powerful deep models (RBMs, Autoencoder, CNN, and RNN), we can review related IoT studies in a new perspective and improve the final system performance. Furthermore, we may propose brand new IoT applications and make a contribution to the "smarter" society.

V. CONCLUSION

In this article, we have investigated how deep learning brings new opportunities to the IoT. It is clear that many IoT applications have been empowered with deep learning tools. Deep models are powerful tools to solve large-scale data analysis problems. We have surveyed the recent research on how to build a deep model for IoT applications by RBM, autoencoder, CNN or RNN. With deep learning, it is unnecessary to take efforts on designing complex features. Deep learning offers us a new perspective to solve traditional problems and help us reveal new insights on the field of IoT. However, how to design an architecture with high accuracy and low resource consumption is still an open issue, and the exploration of this field has not yet come to an end.

REFERENCES

[1] C. Cai, M. Hu, D. Cao, X. Ma, Q. Li, and J. Liu, "Self-deployable indoor localization with acoustic-enabled IoT devices exploiting participatory sensing," *IEEE Internet Things J.*, vol. 6, no. 3, pp. 5297–5311, Jun. 2019, doi: 10.1109/JIOT.2019.2900524.

[2] C. Wang, H. Lin, and H. Jiang, "CANS: Towards congestion-adaptive and small stretch emergency navigation with wireless sensor networks," *IEEE Trans. Mobile Comput.*, vol. 15, no. 5, pp. 1077–1089, May 2016.

[3] M. Hu, W. Liu, K. Peng, X. Ma, W. Cheng, J. Liu, and B. Li, "Joint routing and scheduling for vehicle-assisted multi-drone surveillance," *IEEE Internet Things J.*, vol. 6, no. 2, pp. 1781–1790, Apr. 2019.

[4] M. Hu, W. Liu, J. Lu, R. Fu, K. Peng, X. Ma, and J. Liu, "On the joint design of routing and scheduling for vehicle-assisted multi-UAV inspection," *Future Gener. Comput. Syst.*, vol. 94, pp. 214–223, May 2019.

[5] M. Chen, F. Herrera, and K. Hwang, "Cognitive computing: Architecture, technologies and intelligent applications," *IEEE Access*, vol. 6, pp. 19774–19783, 2018.

[6] T. Huang, L. Lan, X. Fang, P. An, J. Min, and F. Wang, "Promises and challenges of big data computing in health sciences," *Big Data Res.*, vol. 2, no. 1, pp. 2–11, Mar. 2015.

[7] M. Castro-Neto, Y.-S. Jeong, M.-K. Jeong, and L. D. Han, "Online-SVR for short-term traffic flow prediction under typical and atypical traffic conditions," *Expert Syst. Appl.*, vol. 36, no. 3, pp. 6164–6173, Apr. 2009.

[8] S. Mahfouz, F. Mourad-Chehade, P. Honeine, J. Farah, and H. Snoussi, "Target tracking using machine learning and Kalman filter in wireless sensor networks," *IEEE Sensors J.*, vol. 14, no. 10, pp. 3715–3725, Oct. 2014.

[9] F. Wang, Y. Zhu, F. Wang, J. Liu, X. Ma, and X. Fan, "Car4Pac: Last mile parcel delivery through intelligent car trip sharing," *IEEE Trans. Intell. Transp. Syst.*, to be published, doi: 10.1109/TITS.2019.2944134.

[10] B. Yang, J. Yang, J. Xu, and D. Yang, "Area localization algorithm for mobile nodes in wireless sensor networks based on support vector machines," in *Proc. Int. Conf. Mobile Ad-Hoc Sensor Netw. (MSN)*, 2007.

[11] F. Wang, F. Wang, X. Ma, and J. Liu, "Demystifying the crowd intelligence in last mile parcel delivery for smart cities," *IEEE Netw.*, vol. 33, no. 2, pp. 23–29, Mar./Apr. 2019.

[12] X. He, K. Wang, H. Huang, T. Miyazaki, Y. Wang, and S. Guo, "Green resource allocation based on deep reinforcement learning in content-centric IoT," *IEEE Trans. Emerg. Topics Comput.*, to be published.

[13] F. Wang, X. Fan, F. Wang, and J. Liu, "Backup battery analysis and allocation against power outage for cellular base stations," *IEEE Trans. Mobile Comput.*, vol. 18, no. 3, pp. 520–533, Mar. 2018.

[14] H. Li, K. Ota, and M. Dong, "Learning IoT in edge: Deep learning for the Internet of Things with edge computing," *IEEE Netw.*, vol. 32, no. 1, pp. 96–101, Jan./Feb. 2018.

[15] M. A. Alsheikh, S. Lin, D. Niyato, and H. P. Tan, "Machine learning in wireless sensor networks: Algorithms, strategies, and applications," *IEEE Commun. Surveys Tuts.*, vol. 16, no. 4, pp. 1996–2018, 4th Quart., 2014.

[16] R. Miotto, F. Wang, S. Wang, X. Jiang, and J. T. Dudley, "Deep learning for healthcare: Review, opportunities and challenges," *Briefings Bioinf.*, vol. 19, no. 6, pp. 1236–1246, Nov. 2018.

[17] M. Mohammadi, A. Al-Fuqaha, S. Sorour, and M. Guizani, "Deep learning for IoT big data and streaming analytics: A survey," *IEEE Commun. Surveys Tuts.*, vol. 20, no. 4, pp. 2923–2960, 4th Quart., 2018.

[18] J. Wang, Y. Ma, L. Zhang, R. X. Gao, and D. Wu, "Deep learning for smart manufacturing: Methods and applications," *J. Manuf. Syst.*, vol. 48, pp. 144–156, Jul. 2018.

[19] A. Fischer and C. Igel, "An introduction to restricted Boltzmann machines," in *Progress in Pattern Recognition, Image Analysis, Computer Vision, and Applications* (Lecture Notes in Computer Science), vol. 7441, L. Alvarez, M. Mejail, L. Gomez, and J. Jacobo, Eds. Berlin, Germany: Springer, 2012. [Online]. Available: https://link.springer.com/chapter/10.1007/978-3-642-33275-3_2#citeas

[20] R. Salakhutdinov, A. Mnih, and G. Hinton, "Restricted Boltzmann machines for collaborative filtering," in *Proc. 24th Int. Conf. Mach. Learn. (ICML)*, 2007.

[21] G. E. Hinton, S. Osindero, and Y.-W. Teh, "A fast learning algorithm for deep belief nets," *Neural Comput.*, vol. 18, no. 7, pp. 1527–1554, 2006.

[22] D. E. Rumelhart, G. E. Hinton, and R. J. Williams, "Learning representations by back-propagating errors," *Nature*, vol. 323, no. 6088, pp. 533–536, 1986.

[23] G. E. Hinton and R. R. Salakhutdinov, "Reducing the dimensionality of data with neural networks," *Science*, vol. 313, no. 5786, pp. 504–507, 2006.

[24] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. Cambridge, MA, USA: MIT Press, 2016. [Online]. Available: http://www.deeplearningbook.org

[25] H. Lee, A. Battle, R. Raina, and A. Y. Ng, "Efficient sparse coding algorithms," in *Proc. 19th Int. Conf. Neural Inf. Process. Syst. (NIPS)*, 2007.

[26] P. Vincent, H. Larochelle, Y. Bengio, and P.-A. Manzagol, "Extracting and composing robust features with denoising autoencoders," in *Proc. 25th Int. Conf. Mach. Learn. (ICML)*, 2008.

[27] D. H. Hubel and T. N. Wiesel, "Receptive fields and functional architecture of monkey striate cortex," *J. Physiol.*, vol. 195, no. 1, pp. 215–243, 1968.

[28] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," in *Proc. 25th Int. Conf. Neural Inf. Process. Syst. (NIPS)*, 2012.

[29] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," 2014, *arXiv:1409.1556*. [Online]. Available: https://arxiv.org/abs/1409.1556

[30] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, "Going deeper with convolutions," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2015, pp. 1–9.

[31] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 770–778.

[32] Y. LeCun, L. D. Jackel, L. Bottou, C. Cortes, J. S. Denker, H. Drucker, I. Guyon, U. A. Müller, E. Säckinger, P. Simard, and V. Vapnik, "Learning algorithms for classification: A comparison on handwritten digit recognition," in *Neural Networks: The Statistical Mechanics Perspective*. Singapore: World Scientific, 1995, pp. 261–276. [Online]. Available: https://nyuscholars.nyu.edu/en/publications/learning-algorithms-for-classification-a-comparison-on-handwritte)

[33] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Comput.*, vol. 9, no. 8, pp. 1735–1780, Nov. 1997.

[34] J. Chung, C. Gulcehre, K. Cho, and Y. Bengio, "Empirical evaluation of gated recurrent neural networks on sequence modeling," 2014, *arXiv:1412.3555*. [Online]. Available: https://arxiv.org/abs/1412.3555

[35] A. W. Yu, H. Lee, and Q. V. Le, "Learning to skim text," 2017, *arXiv:1704.06877*. [Online]. Available: https://arxiv.org/abs/1704.06877

[36] S. Lai, L. Xu, K. Liu, and J. Zhao, "Recurrent convolutional neural networks for text classification," in *Proc. AAAI Conf. Artif. Intell. (AAAI)*, 2015.

[37] K. Cho, B. van Merriënboer, C. Gulcehre, D. Bahdanau, F. Bougares, H. Schwenk, and Y. Bengio, "Learning phrase representations using RNN encoder-decoder for statistical machine translation," 2014, *arXiv:1406.1078*. [Online]. Available: https://arxiv.org/abs/1406.1078

[38] M. Chen, X. Shi, Y. Zhang, D. Wu, and M. Guizani, "Deep features learning for medical image analysis with convolutional autoencoder neural network," *IEEE Trans. Big Data*, to be published, doi: 10.1109/TBDATA.2017.2717439.

[39] N. Y. Hammerla, S. Halloran, and T. Plötz, "Deep, convolutional, and recurrent models for human activity recognition using wearables," in *Proc. 25th Int. Joint Conf. Artif. Intell. (IJCAI)*, 2016.

[40] Y. Gao, X. Xiang, N. Xiong, B. Huang, H. J. Lee, R. Alrifai, X. Jiang, and Z. Fang, "Human action monitoring for healthcare based on deep learning," *IEEE Access*, vol. 6, pp. 52277–52285, 2018.

[41] J. Zhu, A. Pande, P. Mohapatra, and J. J. Han, "Using deep learning for energy expenditure estimation with wearable sensors," in *Proc. Int. Conf. E-Health Netw., Appl. Services (HealthCom)*, 2015.

[42] A. Y. Hannun, P. Rajpurkar, M. Haghpanahi, G. H. Tison, C. Bourn, M. P. Turakhia, and A. Y. Ng, "Cardiologist-level arrhythmia detection and classification in ambulatory electrocardiograms using a deep neural network," *Nature Med.*, vol. 25, no. 1, p. 65, 2019.

[43] A. Prasoon, K. Petersen, C. Igel, F. Lauze, E. Dam, and M. Nielsen, "Deep feature learning for knee cartilage segmentation using a triplanar convolutional neural network," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent. (MICCAI)*, 2013.

[44] V. Gulshan, L. Peng, M. Coram, M. C. Stumpe, D. Wu, A. Narayanaswamy, S. Venugopalan, K. Widner, T. Madams, J. Cuadros, R. Kim, R. Raman, P. C. Nelson, J. L. Mega, and D. R. Webster, "Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs," *Proc. JAMA*, vol. 316, no. 22, pp. 2402–2410, Dec. 2016.

[45] M. Chen, Y. Zhang, M. Qiu, N. Guizani, and Y. Hao, "SPHA: Smart personal health advisor based on deep analytics," *IEEE Commun. Mag.*, vol. 56, no. 3, pp. 164–169, Mar. 2018.

[46] X. Zeng, K. Cao, and M. Zhang, "MobileDeepPill: A small-footprint mobile deep learning system for recognizing unconstrained pill images," in *Proc. 15th Annu. Int. Conf. Mobile Syst., Appl., Services (MobiSys)*, 2017.

[47] A. R. Lopez, X. Giro-I-Nieto, J. Burdick, and O. Marques, "Skin lesion classification from dermoscopic images using deep learning techniques," in *Proc. 13th IASTED Int. Conf. Biomed. Eng. (BioMed)*, 2017.

[48] W.-J. Chang, L.-B. Chen, C.-H. Hsu, C.-P. Lin, and T.-C. Yang, "A deep learning-based intelligent medicine recognition system for chronic patients," *IEEE Access*, vol. 7, pp. 44441–44458, 2019.

[49] M. Chen, J. Zhou, G. Tao, J. Yang, and L. Hu, "Wearable affective robot," *IEEE Access*, vol. 6, pp. 64766–64776, 2018.

[50] M. Chen, W. Li, Y. Hao, Y. Qian, and I. Humar, "Edge cognitive computing based smart healthcare system," *Future Gener. Comput. Syst.*, vol. 86, pp. 403–411, Sep. 2018.

[51] D. Ravì, C. Wong, B. Lo, and G.-Z. Yang, "A deep learning approach to on-node sensor data analytics for mobile or wearable devices," *IEEE J. Biomed. Health Inform.*, vol. 21, no. 1, pp. 56–64, Jan. 2016.

[52] J. Tang, D. Sun, S. Liu, and J.-L. Gaudiot, "Enabling deep learning on IoT devices," *Computer*, vol. 50, no. 10, pp. 92–96, 2017.

[53] M. Chen, J. Yang, J. Zhou, Y. Hao, J. Zhang, and C. Youn, "5G-smart diabetes: Toward personalized diabetes diagnosis with healthcare big data clouds," *IEEE Commun. Mag.*, vol. 56, no. 4, pp. 16–23, Apr. 2018.

[54] T. Muhammed, R. Mehmood, A. Albeshri, and I. Katib, "UbeHealth: A personalized ubiquitous cloud and edge-enabled networked healthcare system for smart cities," *IEEE Access*, vol. 6, pp. 32258–32285, 2018.

[55] H. Jiang, C. Cai, X. Ma, Y. Yang, and J. Liu, "Smart home based on WiFi sensing: A survey," *IEEE Access*, vol. 6, pp. 13317–13325, 2018.

[56] Y. Gu, Y. Chen, J. Liu, and X. Jiang, "Semi-supervised deep extreme learning machine for Wi-Fi based localization," *Neurocomputing*, vol. 166, pp. 282–293, Oct. 2015.

[57] M. Mohammadi, A. Al-Fuqaha, M. Guizani, and J.-S. Oh, "Semisupervised deep reinforcement learning in support of IoT and smart city services," *IEEE Internet Things J.*, vol. 5, no. 2, pp. 624–635, Apr. 2018.

[58] X. Wang, L. Gao, S. Mao, and S. Pandey, "CSI-based fingerprinting for indoor localization: A deep learning approach," *IEEE Trans. Veh. Technol.*, vol. 66, no. 1, pp. 763–776, Jan. 2017.

[59] J. Wang, X. Zhang, Q. Gao, H. Yue, and H. Wang, "Device-free wireless localization and activity recognition: A deep learning approach," *IEEE Trans. Veh. Technol.*, vol. 66, no. 7, pp. 6258–6267, Jul. 2017.

[60] B. A. Erol, A. Majumdar, J. Lwowski, P. Benavidez, P. Rad, and M. Jamshidi, "Improved deep neural network object tracking system for applications in home robotics," in *Computational Intelligence for Pattern Recognition* (Studies in Computational Intelligence), vol. 777, W. Pedrycz and S. M. Chen, Eds. Cham, Switzerland: Springer, 2018. [Online]. Available: https://link.springer.com/chapter/10.1007/978-3-319-89629-8_14#citeas)

[61] S. Levine, P. Pastor, A. Krizhevsky, J. Ibarz, and D. Quillen, "Learning hand-eye coordination for robotic grasping with deep learning and large-scale data collection," *Int. J. Robot. Res.*, vol. 37, nos. 4–5, pp. 421–436, 2018.

[62] P. Bahl and V. N. Padmanabhan, "RADAR: An in-building RF-based user location and tracking system," in *Proc. 19th Annu. Joint Conf. IEEE Comput. Commun. Soc. (INFOCOM)*, Mar. 2000, pp. 775–784.

[63] A. S. Paul and E. A. Wan, "Wi-Fi based indoor localization and tracking using sigma-point Kalman filtering methods," in *Proc. IEEE/ION Position, Location Navigat. Symp.*, May 2008, pp. 646–659.

[64] C.-L. Wu, L.-C. Fu, and F.-L. Lian, "WLAN location determination in e-home via support vector classification," in *Proc. IEEE Int. Conf. Netw., Sens. Control*, Mar. 2004, pp. 1026–1031.

[65] S. Nikitaki and P. Tsakalides, "Localization in wireless networks via spatial sparsity," in *Proc. 44th Asilomar Conf. Signals, Syst. Comput. (ASILOMAR)*, 2010.

[66] M. Li and H.-J. Lin, "Design and implementation of smart home control systems based on wireless sensor networks and power line communications," *IEEE Trans. Ind. Electron.*, vol. 62, no. 7, pp. 4430–4442, Jul. 2015.

[67] C. Kühnel, T. Westermann, F. Hemmert, S. Kratz, A. Müller, and S. Möller, "I'm home: Defining and evaluating a gesture set for smart-home control," *Int. J. Hum.-Comput. Stud.*, vol. 69, no. 11, pp. 693–704, Oct. 2011.

[68] L. Deng, "Artificial intelligence in the rising wave of deep learning: The historical path and future outlook [perspectives]," *IEEE Signal Process. Mag.*, vol. 35, no. 1, pp. 177–180, Jan. 2018.

[69] S.-H. Baeg, J.-H. Park, J. Koh, K.-W. Park, and M.-H. Baeg, "Building a smart home environment for service robots based on RFID and sensor networks," in *Proc. Int. Conf. Control, Automat. Syst.*, 2007.

[70] W. Huang, G. Song, H. Hong, and K. Xie, "Deep architecture for traffic flow prediction: Deep belief networks with multitask learning," *IEEE Trans. Intell. Transp. Syst.*, vol. 15, no. 5, pp. 2191–2201, Apr. 2014.

[71] Y. Lv, Y. Duan, W. Kang, Z. Li, and F.-Y. Wang, "Traffic flow prediction with big data: A deep learning approach," *IEEE Trans. Intell. Transp. Syst.*, vol. 16, no. 2, pp. 865–873, Apr. 2015.

[72] Z. Zhao, W. Chen, X. Wu, P. C. Y. Chen, and J. Liu, "LSTM network: A deep learning approach for short-term traffic forecast," *IET Intell. Transp. Syst.*, vol. 11, no. 2, pp. 68–75, 2017.

[73] J. Zhang, Y. Zheng, and D. Qi, "Deep spatio-temporal residual networks for citywide crowd flows prediction," in *Proc. AAAI Conf. Artif. Intell. (AAAI)*, 2017.

[74] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: Towards real-time object detection with region proposal networks," in *Proc. 28th Int. Conf. Neural Inf. Process. Syst. (NIPS)*, 2015.

[75] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 779–788.

[76] W. Liu, D. Anguelov, D. Erhan, C. Szegedy, S. Reed, C.-Y. Fu, and A. C. Berg, "SSD: Single shot multibox detector," in *Proc. Eur. Conf. Comput. Vis. (ECCV)*, 2016.

[77] H. Li, Y. Li, and F. Porikli, "DeepTrack: Learning discriminative feature representations online for robust visual tracking," *IEEE Trans. Image Process.*, vol. 25, no. 4, pp. 1834–1848, Apr. 2016.

[78] P. Ondrúška and I. Posner, "Deep tracking: Seeing beyond seeing using recurrent neural networks," in *Proc. AAAI Conf. Artif. Intell. (AAAI)*, 2016.

[79] J. Dequaire, P. Ondrúška, D. Rao, D. Wang, and I. Posner, "Deep tracking in the wild: End-to-end tracking using recurrent neural networks," *Int. J. Robot. Res.*, vol. 37, nos. 4–5, pp. 492–512, 2018.

[80] D. Singh and C. K. Mohan, "Deep spatio-temporal representation for detection of road accidents using stacked autoencoder," *IEEE Trans. Intell. Transp. Syst.*, vol. 20, no. 3, pp. 879–887, Mar. 2019.

[81] B. Wu, F. Iandola, P. H. Jin, and K. Keutzer, "SqueezeDet: Unified, small, low power fully convolutional neural networks for real-time object detection for autonomous driving," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. Workshops (CVPRW)*, Jun. 2017, pp. 129–137.

[82] M. Bojarski, D. Del Testa, D. Dworakowski, B. Firner, B. Flepp, P. Goyal, L. D. Jackel, M. Monfort, U. Müller, J. Zhang, X. Zhang, J. Zhao, and K. Zieba, "End to end learning for self-driving cars," 2016, *arXiv:1604.07316*. [Online]. Available: https://arxiv.org/abs/1604.07316

[83] H. Xu, Y. Gao, F. Yu, and T. Darrell, "End-to-end learning of driving models from large-scale video datasets," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 3530–3538.

[84] N. G. Polson and V. O. Sokolov, "Deep learning for short-term traffic flow prediction," *Transp. Res. C, Emerg. Technol.*, vol. 79, pp. 1–17, Jun. 2017.

[85] L. Li, K. Ota, and M. Dong, "Deep learning for smart industry: Efficient manufacture inspection system with fog computing," *IEEE Trans. Ind. Informat.*, vol. 14, no. 10, pp. 4665–4673, Oct. 2018.

[86] J.-K. Park, B.-K. Kwon, J.-H. Park, and D.-J. Kang, "Machine learning-based imaging system for surface defect inspection," *Int. J. Precis. Eng. Manuf.-Green Technol.*, vol. 3, no. 3, pp. 303–310, Jul. 2016.

[87] J. Wang, J. Zhuang, L. Duan, and W. Cheng, "A multi-scale convolution neural network for featureless fault diagnosis," in *Proc. Int. Symp. Flexible Autom. (ISFA)*, 2016.

[88] L. Wang, X. Zhao, J. Pei, and G. Tang, "Transformer fault diagnosis using continuous sparse autoencoder," *SpringerPlus*, vol. 5, no. 1, p. 448, 2016.

[89] S. Kamijo, Y. Matsushita, K. Ikeuchi, and M. Sakauchi, "Traffic monitoring and accident detection at intersections," *IEEE Trans. Intell. Transp. Syst.*, vol. 1, no. 2, pp. 108–118, Jun. 2000.

[90] M. Chen, Y. Hao, K. Lin, L. Hu, and Z. Yuan, "Label-less learning for traffic control in an edge network," *IEEE Netw.*, vol. 32, no. 6, pp. 8–14, Nov./Dec. 2018.

[91] P. Vincent, H. Larochelle, I. Lajoie, Y. Bengio, and P.-A. Manzagol, "Stacked denoising autoencoders: Learning useful representations in a deep network with a local denoising criterion," *J. Mach. Learn. Res.*, vol. 11, pp. 3371–3408, Dec. 2010.

[92] G. B. Huang, H. Lee, and E. Learned-Miller, "Learning hierarchical representations for face verification with convolutional deep belief networks," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2012, pp. 2518–2525.

[93] J. Lu, V. E. Liong, G. Wang, and P. Moulin, "Joint feature learning for face recognition," *IEEE Trans. Inf. Forensics Security*, vol. 10, no. 7, pp. 1371–1383, Jul. 2015.

[94] S. Zafeiriou, C. Zhang, and Z. Zhang, "A survey on face detection in the wild: Past, present and future," *Comput. Vis. Image Understand.*, vol. 138, pp. 1–24, Sep. 2015.

[95] C. Chen, A. Seff, A. Kornhauser, and J. Xiao, "DeepDriving: Learning affordance for direct perception in autonomous driving," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Dec. 2015, pp. 2722–2730.

[96] P. Sermanet, D. Eigen, X. Zhang, M. Mathieu, R. Fergus, and Y. Lecun, "Overfeat: Integrated recognition, localization and detection using convolutional networks," in *Proc. Int. Conf. Learn. Represent. (ICLR)*, 2014.

[97] B. Huval, T. Wang, S. Tandon, J. Kiske, W. Song, J. Pazhayampallil, M. Andriluka, P. Rajpurkar, T. Migimatsu, R. Cheng-Yue, F. Mujica, A. Coates, and A. Y. Ng, "An empirical evaluation of deep learning on highway driving," 2015, *arXiv:1504.01716*. [Online]. Available: https://arxiv.org/abs/1504.01716

[98] J. Donahue, L. A. Hendricks, S. Guadarrama, M. Rohrbach, S. Venugopalan, T. Darrell, and K. Saenko, "Long-term recurrent convolutional networks for visual recognition and description," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2015, pp. 2625–2634.

[99] F. Pernkopf and P. O'Leary, "Visual inspection of machined metallic high-precision surfaces," *EURASIP J. Adv. Signal Process.*, vol. 2002, no. 7, pp. 667–678, Jul. 2002.

[100] Y. Lei, F. Jia, J. Lin, S. Xing, and S. X. Ding, "An intelligent fault diagnosis method using unsupervised feature learning towards mechanical big data," *IEEE Trans. Ind. Electron.*, vol. 63, no. 5, pp. 3137–3147, May 2016.

[101] D. Li, T. Salonidis, N. V. Desai, and M. C. Chuah, "DeepCham: Collaborative edge-mediated adaptive deep learning for mobile object recognition," in *Proc. IEEE/ACM Symp. Edge Comput. (SEC)*, Oct. 2016, pp. 64–76.

[102] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, "Dropout: A simple way to prevent neural networks from overfitting," *J. Mach. Learn. Res.*, vol. 15, no. 1, pp. 1929–1958, Jan. 2014.

[103] P. Zhao, J. Li, F. Zeng, F. Xiao, C. Wang, and H. Jiang, "ILLIA: Enabling k-anonymity-based privacy preserving against location injection attacks in continuous LBS queries," *IEEE Internet Things J.*, vol. 5, no. 2, pp. 1033–1042, Apr. 2018.

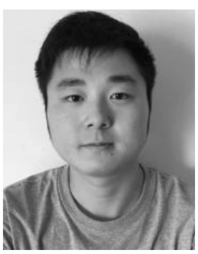
[104] I. Stoica, D. Song, R. A. Popa, D. Patterson, M. W. Mahoney, R. Katz, A. D. Joseph, M. Jordan, J. M. Hellerstein, J. E. Gonzalez, K. Goldberg, A. Ghodsi, D. Culler, and P. Abbeel, "A Berkeley view of systems challenges for AI," 2017, *arXiv:1712.05855*. [Online]. Available: https://arxiv.org/abs/1712.05855

[105] Y. Wei, F. R. Yu, M. Song, and Z. Han, "Joint optimization of caching, computing, and radio resources for fog-enabled IoT using natural actor-critic deep reinforcement learning," *IEEE Internet Things J.*, vol. 6, no. 2, pp. 2061–2073, Apr. 2019.

[106] H. Khelifi, S. Luo, B. Nour, A. Sellami, H. Moungla, S. H. Ahmed, and M. Guizani, "Bringing deep learning at the edge of information-centric Internet of Things," *IEEE Commun. Lett.*, vol. 23, no. 1, pp. 52–55, Jan. 2019.

[107] D. Wang, Y. Peng, X. Ma, W. Ding, H. Jiang, F. Chen, and J. Liu, "Adaptive wireless video streaming based on edge computing: Opportunities and approaches," *IEEE Trans. Services Comput.*, vol. 12, no. 5, pp. 685–697, Sep./Oct. 2019.

XIAOQIANG MA received the B.E. degree from the Huazhong University of Science and Technology, China, in 2010, and the M.Sc. and Ph.D. degrees from Simon Fraser University, Canada, in 2012 and 2015, respectively. His areas of interest are wireless networks, social networks, and cloud computing.



TAI YAO received the B.E. and M.E. degrees from the School of Electronic Information and Communications, Huazhong University of Science and Technology, Wuhan, China, in 2015 and 2019, respectively. He is currently working as a Machine Learning Engineer with Xiaomi Company.

MENGLAN HU received the B.E. degree in electronic and information engineering from the Huazhong University of Science and Technology, China, in 2007, and the Ph.D. degree in electrical and computer engineering from the National University of Singapore, Singapore, in 2012. He is currently an Associate Professor with the School of Electronic Information and Communications, Huazhong University of Science and Technology, China. His research interests include cloud computing, mobile computing, parallel and distributed systems, scheduling and resource management, and wireless networking.

YAN DONG received the B.S. and M.S. degrees from Xidian University, Xi'an, China, and the Ph.D. degree from the Huazhong University of Science and Technology, Wuhan, China, in 2007. She currently works with the School of Electronic Information and Communications, Huazhong University of Science and Technology, as a Professor. Her research interest includes signal processing and coding for high performance wireless networks.

WEI LIU received the B.S. degree in telecommunication engineering and the Ph.D. degree in electronics and information engineering from the Huazhong University of Science and Technology, Wuhan, China, in 1999 and 2004, respectively. He is currently a Professor with the School of Electronic Information and Communications, Huazhong University of Science and Technology. His research interests include the Internet streaming and measurement, and content-centric networks.

FANGXIN WANG received the B.S. degree from the Department of Computer Science of Technology, Beijing University of Post and Telecommunication, Beijing, China, in 2013, and the M.S. degree from the Department of Computer Science and Technology, Beijing, in 2016. He is currently pursuing the Ph.D. degree in the School of Computing Science, Simon Fraser University, Burnaby, BC, Canada. His research interests include the Internet-of-Things, wireless networking, multimedia systems, big data analysis, and machine learning.

JIANGCHUAN LIU received the B.Eng. degree *(cum laude)* from Tsinghua University, Beijing, China, in 1999, and the Ph.D. degree from The Hong Kong University of Science and Technology, in 2003, all in computer science. He is currently a Full Professor (with University Professorship) with the School of Computing Science, Simon Fraser University, British Columbia, Canada.

He is a Steering Committee Member of the IEEE Transactions on Mobile Computing. He is a Fellow of the Canadian Academy of Engineering and an NSERC E.W.R. Steacie Memorial Fellow. He was a co-recipient of the Test of Time Paper Award of ACM Multimedia Best Paper Award, in 2012, ACM TOMCCAP Nicolas D. Georganas Best Paper Award, in 2013, and IEEE INFOCOM, in 2015. He is an Associate Editor of the IEEE/ACM Transactions on Networking, the IEEE Transactions on Big Data, and the IEEE Transactions on Multimedia.