EDITORIAL

Recent advances in deep learning

Xizhao Wang[1] · Yanxia Zhao[1] · Farhad Pourpanah[2]

With the recent advancement in digital technologies, the size of data sets has become too large in which traditional data processing and machine learning techniques are not able to cope with effectively [1, 2]. However, analyzing complex, high dimensional, and noise-contaminated data sets is a huge challenge, and it is crucial to develop novel algorithms that are able to summarize, classify, extract important information and convert them into an understandable form [3–5]. To undertake these problems, deep learning (DL) models have shown outstanding performances in the recent decade.

Deep learning (DL) has revolutionized the future of artificial intelligence (AI). It has solved many complex problems that existed in the AI community for many years. In fact, DL models are deeper variants of artificial neural networks (ANNs) with multiple layers, whether linear or non-linear. Each layer is connected to its lower and upper layers through different weights. The capability of DL models in learning hierarchical features from various types of data, e.g., numerical, image, text and audio, makes them powerful in solving recognition, regression, semi- supervised and unsupervised problems [6–8].

In recent years, various deep architectures with different learning paradigm are quickly introduced to develop machines that can perform similar to human or even better in different domains of application such as medical diagnosis, self-driving cars, natural language and image processing, and predictive forecasting [9]. To show some recent advances of deep learning to some extent, we select 14

papers from the articles accepted in this journal to organize this issue. Focusing on recent developments in DL architectures and their applications, we classify the articles in this issue into four categories: (1) deep architectures and conventional neural networks, (2) incremental learning, (3) recurrent neural networks, and (4) generative models and adversarial examples. In the following, we give a brief summary of each category and then individually introduce related articles.

1 Category 1: deep architectures and conventional neural networks

Deep neural network (DNN) [10] is one of the most common DL models that contains multiple layers of linear and non-linear operations. DNN is the extension of standard neural network with multiple hidden layers, which allows the model to learn more complex representations of the input data. In addition, convolutional neural network (CNN) is a variant of DNNs, which is inspired by the visual cortex of animals [11]. CNN usually contains three types of layers, including convolution, pooling, and fully connected layers. The convolution and pooling layers are added in the lower levels. The convolution layers generate a set of linear activations, which is followed by non-linear functions. In fact, the convolution layers apply some filters to reduce complexity of the input data [12]. Then, the pooling layers are used for down-sampling of the filtered results. The pooling layers manage to reduce the size of the activation maps by transferring them into a smaller matrix [13]. Therefore, pooling solves the over-fitting problem by reducing complexity [14]. The fully connected layers are located after the convolution and pooling layers, in order to learn more abstract representations of the input data. In the last layer, a loss function, e.g., a soft-max classifier, is used to map the input data to its corresponding output. CNN-based models have shown outstanding results in the areas of image processing and computer vision. This category contains four articles.

✉ Xizhao Wang
xizhaowang@ieee.org

Yanxia Zhao
confidenceaaa@126.com

Farhad Pourpanah
farhad@szu.edu.cn

1 College of Management, Hebei University, Baoding 071002, Hebei, China

2 College of Mathematics and Statistics, Shenzhen University, Shenzhen 518060, Guangdong, China

The paper "Combination of loss functions for deep text classification" authored by Hamideh Hajiabadi, Diego Molla-Aliod, Reza Monsefi and Hadi Sadoghi Yazdi, considers ensemble methods at the level of the objective function of a deep neural network. This paper proposed a novel objective function that is a linear combination of single losses and integrate the proposed objective function into a deep neural network. By doing so, the weights associated with the linear combination of losses are learned by back propagation during the training stage. The impact of the proposed ensemble loss function is studied on the state-of-the-art convolutional neural networks for text classification.

In the paper "A deep neural network-based recommendation algorithm using user and item basic data" authored by Jian-Wu Bi, Yang Liu and Zhi-Ping Fan, a new recommendation algorithm based on deep neural networks is proposed. The main idea of the algorithm is to build a regression model for predicting user ratings based on deep neural networks. To this end, based on user data and item data, a user feature matrix and an item feature matrix are respectively constructed using the four types of neural network layers, i.e., embedding layer (EL), convolution layer (CL), pooling layer (PL) and fully connected layer (FCL). Then, based on the obtained matrixes, a user-item feature matrix is further constructed using a FCL. On this basis, a regression model for predicting user ratings is trained to generate recommendation list.

The paper "A discriminative deep association learning for facial expression recognition" authored by Xing Jin, Wenyun Sun and Zhong Jin, proposed a novel discriminative deep association learning (DDAL) framework for facial expression recognition. In this work, the unlabeled data is used to train the DNNs with the labeled data simultaneously, in a multi-loss deep network based on association learning. In addition, the discrimination loss is utilized to ensure intraclass clustering and inter-class centers separating.

In the paper "A technical view on neural architecture search" authored by Yi-Qi Hu and Yang Yu, a review of recent advances in neural architecture search (NAS) from a technical point of view is provided. This paper drew a whole picture of NAS for readers including problem definition, basic search framework, key technique towards practice and promising future directions.

2 Category 2: incremental learning

Incremental learning refers to the condition of continuous model adaptation based on a constantly arriving input samples [15–17]. Unlike machine learning techniques with batch learning procedure that have to re-execute an iterative training procedure using both old and new samples, incremental learning techniques require to learn only new samples without re-learning preciously learned samples [18, 19]. Besides, incremental learning techniques are useful for training complex structure of DL models when the training samples are provided over time [20, 21]. This category contains two articles.

The paper "Cross-modal learning for material perception using deep extreme learning ma- chine", authored by Wendong Zheng, Huaping Liu, Bowen Wang and Fuchun Sun, proposed a visual-tactile cross-modal retrieval framework to convey tactile information of surface material for perceptual estimation. In this paper, tactile information of a new unknown surface material is used to retrieve perceptually similar surface from an available surface visual sample set. Specifically, a deep cross-modal correlation learning method, which incorporates the high-level nonlinear representation of deep extreme learning machine and class-paired correlation learning of cluster canonical correlation analysis, is developed.

The paper "DeepCascade-WR: a cascading deep architecture based on weak results for time series prediction" authored by Chunyang Zhang, Qun Dai and Gang Song, considers the real- world time series predictions (TSPs) tasks. In this work, a cascading deep architecture based on weak results (DeepCascade-WR) is established, which possesses deep models marked capability of feature representation learning based on complex data. DeepCascade-WR possesses online learning ability and effectively avoids the retraining problem, owing to the property of OS-ELM. In addition, DeepCascade-WR naturally inherits some valuable virtues from ELM, including faster training speed, better generalization ability and the avoidance of being fallen into local optima.

3 Category 3: recurrent neural networks

Recurrent neural networks (RNNs) [22] have the deepest structures among the DL algorithms, which are able to map sequential input data to their output [23]. Unlike traditional DNNs, the nodes in each RNN layer are connected to each other. This self-connection enables RNNs to memorize information over time from a sequence of data. The long-short term memory (LSTM) [24] and gated recurrent units (GRU) [25] are two improved models of RNNs. Although RNNs are powerful, it is difficult to train a long-range sequence of data due to vanishing or exploding gradient problem [26]. To solve this issue, LSTM and GRU use gate units to decide what information to keep or remove from the previous state. RNN-based models have been widely applied to handle sequential learning problems. This category contains five articles.

The paper "DeepSite: bidirectional LSTM and CNN models for predicting DNAprotein binding" authored by

Yongqing Zhang, Shaojie Qiao, Shengjie Ji and Yizhou Li, considers the prediction of DNA–protein binding sites in DNA sequence using DL methods. In this paper, Deep-Site, which is the bidirectional long short-term memory (BLSTM) and CNN, is employed to capture the long-term dependencies between the sequence motifs in DNA.

The paper "Single image rain streaks removal: a review and an exploration" authored by Hong Wang, Qi Xie, Yichen Wu, Qian Zhao and Deyu Meng, provided a detailed review of single-image-based rain removal techniques in recent years. These techniques are categorized into: early filter-based, conventional prior-based, and recent deep learning-based approaches. In addition, inspired by the rationality of DL-based methods and insightful characteristics underlying rain shapes, a specific coarse-to-fine de-raining network architecture is built. This architecture is able to deliver the rain structures and progressively removes rain streaks from the input image, accordingly.

The paper "Learning deep hierarchical and temporal recurrent neural networks with residual learning" authored by Tehseen Zia, Assad Abbas, Usman Habib and Muhammad Sajid Khan, studies deep hierarchical and temporal structures in RNNs. The goal is to prove that approximating identity mapping is crucial for optimizing both hierarchical and temporal structures. In this regard, a framework, called hierarchical and temporal residual RNNs, is proposed to learn RNNs by approximating identity mappings across hierarchical and temporal structures.

The paper "Weighted multi-deep ranking supervised hashing for efficient image retrieval" authored by Jiayong Li, Wing W. Y. Ng, Xing Tian, Sam Kwong and Hui Wang, focuses on deep hashing networks for large-scale image retrieval. This paper proposed a weighted multi-deep ranking supervised hashing (WMDRH), which employs multiple weighted deep hash tables, to improve precision/recall without increasing space usage. A loss function that contains two terms: (1) the ranking pairwise loss and (2) the classification loss, is used to generate hash codes. The former one ensures to generate discriminative hash codes by penalizing more for the (dis)similar image pairs with (small)large Hamming distances, and the classification loss guarantees the hash codes to be effective for category prediction. Besides, multiple hash tables are integrated by assigning the appropriate weight to each table according to its mean average precision (MAP) score for image retrieval.

The paper "Pothole detection using location-aware convolutional neural networks" authored by Hanshen Chen, Minghai Yao and Qinlong Gu, proposed a new method based on location- aware convolutional neural networks to detect pothole in road images. The proposed method consists of two subnetworks: the first one employs a high-recall network model to find as many candidate regions as possible, and the second one performs classification on the candidates on which the network is expected to focus.

4 Category 4: generative models and adversarial examples

Generative models aim to generate new samples with some variations through learning distribution of the training samples [27]. Variational autoencoders (VAE) [28] and generative adversarial networks (GAN) [29] are two prominent members of generative models. DL models usually require large amount of labeled samples to learn their parameters. However, obtaining sufficient labeled samples in many practical applications is difficult and computationally expensive. To alleviate this problem, generative models can be used [30]. They can be used to solve recognition, semi-supervised learning, unsupervised feature learning, denoising tasks. Despite the great successes of DL models in solving many real-world problems, they can be easily fooled by adversarial examples [31]. This issue raises concerns in many fields such as safety or autonomous vehicles. Thus it is crucial to study the effects of adversarial examples on the performance of DL models. This category contains three articles.

In the paper "An adversarial non-volume preserving flow model with Boltzmann priors" authored by Jian Zhang, Shifei Ding and Weikuan Jia, an adversarial non-volume preserving flow model with Boltzmann priors (ANVP) for modeling complex high-dimensional densities is proposed. ANVP introduced an adversarial regularizer into the loss function to penalize the condition that places a high probability in regions where the training data distribution has a low density to generate sharper images.

The paper "Emotion recognition using multimodal deep learning in multiple psychophysiological signals and video" authored by Wang Zhongmin, Zhou Xiaoxiao, Wang Wenlang and Liang Chen, proposed an DL based approach to train several specialist networks to fuse the features of individual modalities. This approach includes a multimodal deep belief network (MDBN) and two bimodal deep belief network (BDBN). The MDBN is used to optimize and fuse unified psychophysiological features derived from the features of multiple psychophysiological signals, one DBBN to focus on representative visual features among the features of a video stream, and another DBBN to focus on the high multimodal features in the unified features obtained from two modalities.

The paper "Robustness to Adversarial Examples can be Improved with Overfitting" authored by Oscar Deniz, Noelia Vallez, Jesus Salido and Gloria Bueno, studies the effects of adversarial examples on the performance of DL methods. This paper, firstly, argued that the error in ad- versarial examples is caused by high bias, i.e. by regularization that

has local negative effects, and then supported this idea by experiments in which the robustness to adversarial examples is measured with respect to the level of fitting to training samples.

In summary, this issue shows some recent advances in DL from a new angle to some extent. It includes fourteen articles belonging to four categories, among which four belong to the scope of deep architectures and conventional neural networks, two belong to the area of incremental learning, five belong to the scope of recurrent neural networks, and three belong to the field of generative models and adversarial examples. It aims to provide readers with some useful guidelines to know the recent developments in algorithm and application of DL, and to give a collection of DL articles for readers convenient to reference.

References

1. Wang X, Joshua HZ (2015) Uncertainty in learning from big data. Fuzzy Sets Syst 258:1–4
2. Rezvani S, Wang X, Pourpanah F (2019) Intuitionistic fuzzy twin support vector machines. IEEE Trans Fuzzy Syst 27(11):2140–2151
3. Wang Z, Wang X (2018) A deep stochastic weight assignment network and its application to chess playing. J Parallel Distrib Comput 117:205–211
4. Sherkatghanad Z, Akhondzadeh M, Salari S, Zomorodi-Moghadam M, Abdar M, Acharya UR, Khosrowabadi R, Salari V (2019) Automated detection of autism spectrum disorder using a convolutional neural network. Front Neurosci 13:1325
5. Pourpanah F, Lim CP, Wang X, Tan CJ, Seera M, Shi Y (2019) A hybrid model of fuzzy minmax and brain storm optimization for feature selection and data classification. Neurocomputing 333:440–451
6. LeCun Y, Bengio Y, Hinton G (2015) Deep learning. Nature 521(7553):36
7. Wang X, Zhang T, Wang R (2017) Noniterative deep learning: incorporating restricted boltzmann machine into multilayer random weight neural networks. IEEE Trans Syst Man Cybern Syst 49(7):1299–1308
8. Korsuk S, Ahmed RSE, Yee-Wah T, David SRJ, Cree IA, Rajpoot NM (2016) Locality sensitive deep learning for detection and classification of nuclei in routine colon cancer histology images. IEEE Trans Med Imaging 35(5):1196–1206
9. Sengupta S, Basak S, Saikia P, Paul S, Tsalavoutis V, Atiah F, Ravi V, Peters A (2020) A review of deep learning with special emphasis on architectures, applications and recent trends. Knowl-Based Syst https://doi.org/10.1016/j.knosys.2020.105596
10. Saxe AM, McClelland JL, Ganguli S (2013) Exact solutions to the nonlinear dynamics of learning in deep linear neural networks. arXiv:1312.6120
11. Hubel DH, Wiesel TN (1968) Receptive fields and functional architecture of monkey striate cortex. J Physiol 195(1):215–243
12. Hinton GE, Srivastava N, Krizhevsky A, Sutskever I, Salakhutdinov RR (2012) Improving neural networks by preventing co-adaptation of feature detectors. arXiv:1207.0580
13. Scherer D, Mller A, Behnke S (2010) Evaluation of pooling operations in convolutional architectures for object recognition. In: International conference on artificial neural networks, pp 92–101
14. Nguyen A, Yosinski J, Clune J (2015) Deep neural networks are easily fooled: High confidence predictions for unrecognizable images. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 427–436
15. Pourpanah F, Wang R, Lim CP, Wang X, Seera M, Tan CJ (2019) An improved fuzzy ARTMAP and Q-learning agent model for pattern classification. Neurocomputing 359:139–152
16. Silver DL (2011) Machine lifelong learning: challenges and benefits for artificial general intelligence. In: International conference on artificial general intelligence, pp 370–375
17. Pourpanah F, Lim CP, Hao Q (2019) A reinforced fuzzy ARTMAP model for data classification. Int J Mach Learn Cybernet 10(7):1643–1655
18. Jain LC, Seera M, Lim CP, Balasubramaniam P (2014) A review of online learning in supervised neural networks. Neural Comput Appl 25:491–509
19. Pourpanah F, Zhang B, Ma R, Hao Q (2018) Non-intrusive human motion recognition using distributed sparse sensors and the genetic algorithm based neural network. IEEE Sensors, pp 1–4
20. Gepperth A, Hammer B (2016) Incremental learning algorithms and applications. In: European symposium on artificial neural networks
21. Sarwar SS, Ankit A, Roy K (2020) Incremental learning in deep convolutional neural networks using partial network sharing. IEEE Access 8:4615–4628
22. Pascanu R, Gulcehre C, Cho K, Bengio Y (2013) How to construct deep recurrent neural networks. arXiv:1312.6026
23. Rumelhart DE, Hinton GE, Williams RJ (1986) Learning representations by back-propagating errors. Nature 323(6088):533–536
24. Hochreiter S, Schmidhuber J (1997) Long short-term memory. Neural Comput 9(8):1735–1780
25. Chung J, Gulcehre C, Cho K, Bengio Y (2014) Empirical evaluation of gated recurrent neural networks on sequence modeling. arXiv:1412.3555
26. Dang A, Vu TH, Wang JC (2017) A survey of deep learning for polyphonic sound event detection. In: International conference on orange technologies, pp 75–78
27. Koller D, Friedman N (2009) Probabilistic graphical models: principles and techniques. MIT Press, Cambridge
28. Kingma DP, Welling M (2014) Auto-encoding variational Bayes. In: International conference on learning representations
29. Goodfellow I, Abadie JP, Mirza M, Xu B, Farley DW, Ozair S, Courville A, Bengio Y (2014) Generative adversarial nets. In: Neural Information Processing Systems, pp 2672–2680
30. Zhijian O (2018) A review of learning with deep generative models from perspective of graphical modeling. arXiv:1808.01630
31. Szegedy C, Zaremba W, Sutskever I, Bruna J, Erhan D, Goodfellow I, Fergus R (2013) Intriguing properties of neural networks. arXiv:1312.6199