POLITECNICO DI TORINO
Repository ISTITUZIONALE

Special Issue on Advances in Deep Learning

Original
Special Issue on Advances in Deep Learning / Gragnaniello, D., Bottino, A., Cumani, S., Kim, W.. - In: APPLIED SCIENCES. - ISSN 2076-3417. - ELETTRONICO. - 10:9(2020), pp. 3172-3177. [10.3390/app10093172]

Availability:
This version is available at: 11583/2823632 since: 2020-05-13T12:22:46Z

Publisher:
MDPI

Published
DOI:10.3390/app10093172

Terms of use:

This article is made available under terms and conditions as specified in the corresponding bibliographic description in the repository

Publisher copyright

(Article begins on next page)

17 November 2026

applied sciences

MDPI

Editorial

Special Issue on Advances in Deep Learning

Diego Gragnaniello [1,*], Andrea Bottino [2], Sandro Cumani [2] and Wonjoon Kim [3]

[1] Department of Electrical Engineering and Information Technology, Università degli Studi di Napoli Federico II, 80125 Napoli, Italy
[2] Department of Control and Computer Engineering, Politecnico di Torino, 10129 Torino, Italy; andrea.bottino@polito.it (A.B.); sandro.cumani@polito.it (S.C.)
[3] Department of Industrial and Management Engineering, Sungkyul University, Anyang 14907, Korea; wjkim@sungkyul.ac.kr
* Correspondence: diego.gragnaniello@unina.it

Received: 4 April 2020; Accepted: 25 April 2020; Published: 2 May 2020

1. Introduction

Nowadays, deep learning is the fastest growing research field in machine learning and has a tremendous impact on a plethora of daily life applications, ranging from security and surveillance to autonomous driving, automatic indexing and retrieval of media content, text analysis, speech recognition, automatic translation, and many others. The lightning fast progress of the research in deep learning is testified by the success of this special issue, which witnessed the submission of promising works in several fields like computer vision, signal processing, and natural language processing. The heterogeneity of the proposed approaches also characterizes these works, which span discriminative and generative methods, adversarial perturbations, and reinforcement learning. For better readability, the rest of this editorial summarizes the works published in this special issue on a per-theme basis.

2. Content

Computer vision. Most of the works accepted to this special issue deal with computer vision applications, in general, and with image classification and object detection, in particular, with authors putting considerable effort in proposing novel and state-of-the-art classification techniques. Various contributions tackled (with interesting results) challenging issues like object detection and person re-identification in the wild. For instance, both the authors of [1,2] propose deep learning approaches aimed at recognizing body parts. In [1], the semantic relationship between body segments is automatically learned from videos and used to improve pedestrian detection when occlusions are present in the video. On the same path, the authors of [2] use the detected body parts for improving person re-identification. The method combines four CNN branches: one for encoding the whole body appearance and the other for extracting robust embeddings from three image patches, corresponding to different body parts (head, torso and lower-body) and segmented through U-net [3]. Then, the final person identification layer leverages a classifier trained on the fused outputs of these CNNs. The authors of [4] approach the identification of anomalous events for video surveillance tasks by using a 3D CNN that is merely trained on "normal" events. The resulting model is, thus, a robust outlier detector capable of adapting to rare events.

Infrared imagery is also gaining great interest due to its relevance in critical applications. Inspired by Faster R-CNN [5], object detection in infrared streetscape images is addressed in [6] by exploiting both fine- and coarse-grained image features. Among applications in real-world scenarios, text detection is also gaining attention for the automotive industry. In [7], the authors focus on text detection from Google Street View images. By combining a region proposal (obtained through an

attention-based CNN) with a semantic segmentation (which results from a fully connected network), the proposed technique is capable of reliably extracting text from multiple image regions in parallel.

Industry is another area of large interest for the research community. A specific problem to address in this context is the development of robust and reliable image-based diagnostic tools. As an example, the authors of [8] present an investigation on the use of CNNs for heated metal attributes such as heating temperature or duration, cooling mode, and relative humidity, with particular attention to low computational complexity models that allow deployment on mobile devices. Remote sensing images is another scenario where classical computer vision applications like object detection, scene classification, and scene retrieval are relevant. In this regard, the survey in [9] collects the most recent contributions in these specific application fields.

Other works accepted to this special issue address more general questions related to classification performance of deep neural networks. As an example, the authors of [10,11] present two different approaches to reduce the complexity of CNNs, thus allowing a lighter training phase and a faster inference. In particular, the authors of [10] propose a CNN compression method that exploits kernel density estimation to perform a fast 4-bit quantization of the network weights without impairing classification performance. On the contrary, the authors of [11] introduce the Layer Selective Learning framework, which aims to improve the distillation of a very deep and well-trained network into a shallower (and much faster) one. Furthermore, the work shows that the proposed distillation approach also improves the generalization capability of deep classifiers when tackling new domains where a limited amount of labeled data is available. Finally, the work in [12] addresses the difficulties in fine-tuning the hyperparameters of the architecture to be trained. The authors propose the use of an evolutionary approach to automatically set (for each specific task or dataset) a variety of parameters, like the number of layers and their size.

Generative models. A vast research field in deep learning regards generative models, as witnessed by the ability of recent approaches to learn data distribution and generate new complex samples, like human faces that are barely distinguishable from the original one. An application of such generative models is the inpainting of occluded or missing image regions. As an example, the authors of [13] present a method that transforms the depth image inpainting into the dual problem of image denoising, which is then solved with the help of learned CNN denoiser priors. A novel generative model (Encapsulated Variational Autoencoders) is presented in [14]. Autoencoders can learn, thanks to their bottleneck-shaped architecture, general and effective latent representations for generative tasks. The proposed approach stacks two variational autoencoders controlling their mutual dependence in order to improve both the generative and the discriminative power of the whole architecture. Generative Adversarial Networks (GANs) are another approach to (deep) generative modeling. The authors of [15] present an example of their application to the image-to-image translation, where a generator takes as input an image of the source domain and outputs a related image belonging to the target domain. The model is trained with an adversarial loss that discriminates between identical and generated image pairs.

Another extensive application area for generative models is domain adaptation (DA). DA has rapidly gained attention from the scientific community, thanks to its contribution to approaching real cases where a sufficient amount of training data is not available. As shown in [16], GANs and autoencoders can be combined to provide an unsupervised DA method for image classification. However, GAN-based DA requires facing several challenges such as deciding when to stop training, how to properly evaluate the training performance when building a validation set is not possible, and how to guarantee the reliability of the model output in case of significant domain shifts. To soften these issues, the authors of [17] introduce novel confidence metrics to predict the generalization ability of the model and thus stop its training without the need for a separate validation set.

Another use of generative models is the development of powerful adversarial attacks, where adding purposefully tiny perturbations to a clean input can lead the classifier to output the label of any desired class. In [18], an efficient method to improve the attack success rate in the most challenging scenario, i.e., when the attacker has low or no knowledge about the model to attack,

is presented. The proposed approach exploits data augmentation to train an ensemble of substitute models that are used to craft adversarial examples. Given the relevance of the threat of such adversarial attacks, the survey in [19] tries to resume recent advances in this rapidly growing subject area, from both the defensive and offensive points of view.

One-dimensional signals. Besides computer vision applications, deep learning is widespread in one-dimensional signal processing. As an example, in [20], audio speech is enhanced through deep learning. This work presents an interesting denoising approach based on the disentangling of the representation of the useful signal and the noise. These two components are then discriminated in the latent space by a neural network trained through an adversarial paradigm. Another relevant application related to audio processing is recognizing emotions in human speech. This information can then be used in various domains, from customer care to security. The authors of [21] tackle this problem by feeding a Long Short-Term Memory (LSTM) network with spectrograms and handcrafted features capturing voice timber. The final emotion classification is then obtained through a Support Vector Machine. Similar approaches are adopted to address signal classification and information extraction in different domains. As an example, in [22], CNN and particle swarm optimization are jointly employed to predict large-scale wind power, while in [23], a self-attention CNN is used to classify the heartbeat signal of hatching eggs in order to tell apart dead from alive ones in commercial poultry breeding. In [24], the authors present an attention-based deep learning approach to automatic signal modulation recognition, which is robust against transmission errors. Finally, in [25,26], the classification of vibration signals is carried out through recurrent neural networks. In the first work [25], bearing fault diagnosis is performed to prevent accidents and reduce production costs, while the second paper [26] is concerned with surface roughness prediction for milling process quality assessment.

Text. Deep learning approaches also appear in many applications related to text processing, in both structured (e.g., web pages and source code) and unstructured domains (e.g., Natural Language Processing (NLP)). This special issue presents two approaches to structured text processing, both leveraging recurrent models but with very different aims. In [27], track's metadata and user's behavioral data are jointly used to train an LSTM, which learns user preferences and can improve personalized music recommendation. Instead, the authors of [28] try to predict the risk of chronic disease in patients analyzing the data collected from a national survey designed to assess the health and nutritional status of adults and children in Korea. In particular, they built their approach around Char-RNN [29], a deep recurrent network that efficiently models language by also handling missing data in input sequences.

Concerning unstructured text (NLP), various applications try to reach a data representation at a high level of abstraction to help correctly extract the information in a way that is independent from the words used to express it. In this regard, there are two works accepted in this special issue that propose sequence-to-sequence based summarization approaches. The first [30] promotes diversity in the final summary through an attention-based approach. The second [31] proposes a generative method (based on CNNs and equipped with both an attention mechanism and a module for identifying rare/unseen words) that can spot key sentences and capture the hierarchical structure of textual documents. Other NLP approaches are discussed in [32], which adopts Self Organizing Maps and Bag of Words to measure text similarity, and in [33], which investigates the use of capsule networks (initially proposed for computer vision applications [34]) for sentence classification and sentiment analysis. The work in [35] tackles a different problem, i.e., providing human-understandable justifications of the classification output for sentiment analysis from unstructured text. The approach exploits a Discretized Interpretable Multi-Layer Perceptron architecture to transform the last connected layer of a CNN into a set of rules capable of explaining with high fidelity the classification decision. Finally, the work in [36] uses deep learning to process multiple sources of heterogeneous information jointly. In particular, the authors propose a framework that separately processes text (through an LSTM) and images attached to emails (through a CNN) and then fuses their classification probabilities to increase the spam filtering robustness.

Reinforcement learning. Finally, a recent trend in research is adopting the reinforcement learning paradigm in a variety of real-world applications ranging from robotics to automotive, gaming, personalized recommendations, and advertising. This technique enables agents to learn the best possible action in the environment by assigning penalties and rewards to all possible choices. In this special issue, we have two examples of reinforcement learning applications: the dredger control system presented in [37], which is trained on historical data acquired during human-controlled dredging operations, and an intelligent controller of the water level in a multi-input multi-output communicating two-tank system [38].

3. Conclusions

The diverseness of the works published in this Special Issue confirms the widespread of deep learning in several aspects of our life. In the scientific community, new learning strategies are regularly proposed while, concurrently, more robust techniques are presented in very challenging scenarios. This expansion is catalyzed by the availability of more data and more powerful computational resources. In this special issue, recent deep learning strategies exploited in very different applications have been collected with the hope of inspiring cross-contamination of ideas. However, a constant bibliography update is necessary to keep track of the fast evolution of this research field.

Author Contributions: Contribution is equally divided among all authors. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Acknowledgments: This special issue would not be possible without the contributions of various talented authors, professional and hardworking reviewers, and committed editorial team of Applied Sciences. We thank all authors, and we hope that, no matter what the final decisions of the submitted manuscripts in our special issue were, the feedback and suggestions from the reviewers and editors helped the authors to improve their papers. We would like to take this opportunity to record our sincere gratefulness to all reviewers. Without their dedication, our special issue would not have succeeded. Finally, we place on record our gratitude to the editorial team of Applied Sciences and special thanks to Daria Shi, Managing Editor, from MDPI Branch Office, Beijing.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Gu, J.; Lan, C.; Chen, W.; Han, H. Joint Pedestrian and Body Part Detection via Semantic Relationship Learning. *Appl. Sci.* **2019**, *9*, 752. doi:10.3390/app9040752. [CrossRef]
2. Gao, H.; Chen, S.; Zhang, Z. Parts Semantic Segmentation Aware Representation Learning for Person Re-Identification. *Appl. Sci.* **2019**, *9*, 1239. doi:10.3390/app9061239. [CrossRef]
3. Ronneberger, O.; Fischer, P.; Brox, T. U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*; Springer: Berlin/Heidelberg, Germany, 2015; pp. 234–241.
4. Bouindour, S.; Snoussi, H.; Hittawe, M.M.; Tazi, N.; Wang, T. An On-Line and Adaptive Method for Detecting Abnormal Events in Videos Using Spatio-Temporal ConvNet. *Appl. Sci.* **2019**, *9*, 757. doi:10.3390/app9040757. [CrossRef]
5. Ren, S.; He, K.; Girshick, R.; Sun, J. Faster r-cnn: Towards real-time object detection with region proposal networks. *Adv. Neural Inf. Process. Syst.* **2015**, *28*, 91–99. [CrossRef] [PubMed]
6. Qu, H.; Zhang, L.; Wu, X.; He, X.; Hu, X.; Wen, X. Multiscale Object Detection in Infrared Streetscape Images Based on Deep Learning and Instance Level Data Augmentation. *Appl. Sci.* **2019**, *9*, 565. doi:10.3390/app9030565. [CrossRef]
7. Qin, H.; Zhang, H.; Wang, H.; Yan, Y.; Zhang, M.; Zhao, W. An Algorithm for Scene Text Detection Using Multibox and Semantic Segmentation. *Appl. Sci.* **2019**, *9*, 1054. doi:10.3390/app9061054. [CrossRef]
8. Yin, H.; Mao, K.; Zhao, J.; Chang, H.; E, D.; Tan, Z. Heated Metal Mark Attribute Recognition Based on Compressed CNNs Model. *Appl. Sci.* **2019**, *9*, 1955. doi:10.3390/app9091955. [CrossRef]

9. Gu, Y.; Wang, Y.; Li, Y. A Survey on Deep Learning-Driven Remote Sensing Image Scene Understanding: Scene Classification, Scene Retrieval and Scene-Guided Object Detection. *Appl. Sci.* **2019**, *9*, 2110. doi:10.3390/app9102110. [CrossRef]
10. Seo, S.; Kim, J. Efficient Weights Quantization of Convolutional Neural Networks Using Kernel Density Estimation based Non-uniform Quantizer. *Appl. Sci.* **2019**, *9*, 2559. doi:10.3390/app9122559. [CrossRef]
11. Li, H.T.; Lin, S.C.; Chen, C.Y.; Chiang, C.K. Layer-Level Knowledge Distillation for Deep Neural Network Learning. *Appl. Sci.* **2019**, *9*, 1966. doi:10.3390/app9101966. [CrossRef]
12. Park, K.m.; Shin, D.; Chi, S.D. Variable Chromosome Genetic Algorithm for Structure Learning in Neural Networks to Imitate Human Brain. *Appl. Sci.* **2019**, *9*, 3176. doi:10.3390/app9153176. [CrossRef]
13. Li, Z.; Wu, J. Learning Deep CNN Denoiser Priors for Depth Image Inpainting. *Appl. Sci.* **2019**, *9*, 1103. doi:10.3390/app9061103. [CrossRef]
14. Bai, W.; Quan, C.; Luo, Z.W. Improving Generative and Discriminative Modelling Performance by Implementing Learning Constraints in Encapsulated Variational Autoencoders. *Appl. Sci.* **2019**, *9*, 2551. doi:10.3390/app9122551. [CrossRef]
15. Sung, T.L.; Lee, H.J. Image-to-Image Translation Using Identical-Pair Adversarial Networks. *Appl. Sci.* **2019**, *9*, 2668. doi:10.3390/app9132668. [CrossRef]
16. Wang, X.; Wang, X. Unsupervised Domain Adaptation with Coupled Generative Adversarial Autoencoders. *Appl. Sci.* **2018**, *8*, 2529. doi:10.3390/app8122529. [CrossRef]
17. Bonechi, S.; Andreini, P.; Bianchini, M.; Pai, A.; Scarselli, F. Confidence Measures for Deep Learning in Domain Adaptation. *Appl. Sci.* **2019**, *9*, 2192. doi:10.3390/app9112192. [CrossRef]
18. Gao, X.; Tan, Y.a.; Jiang, H.; Zhang, Q.; Kuang, X. Boosting Targeted Black-Box Attacks via Ensemble Substitute Training and Linear Augmentation. *Appl. Sci.* **2019**, *9*, 2286. doi:10.3390/app9112286. [CrossRef]
19. Qiu, S.; Liu, Q.; Zhou, S.; Wu, C. Review of Artificial Intelligence Adversarial Attack and Defense Technologies. *Appl. Sci.* **2019**, *9*, 909. doi:10.3390/app9050909. [CrossRef]
20. Bae, S.H.; Choi, I.; Kim, N.S. Disentangled Feature Learning for Noise-Invariant Speech Enhancement. *Appl. Sci.* **2019**, *9*, 2289. doi:10.3390/app9112289. [CrossRef]
21. Tursunov, A.; Kwon, S.; Pang, H.S. Discriminating Emotions in the Valence Dimension from Speech Using Timbre Features. *Appl. Sci.* **2019**, *9*, 2470. doi:10.3390/app9122470. [CrossRef]
22. Yang, X.; Zhang, Y.; Yang, Y.; Lv, W. Deterministic and Probabilistic Wind Power Forecasting Based on Bi-Level Convolutional Neural Network and Particle Swarm Optimization. *Appl. Sci.* **2019**, *9*, 1794. doi:10.3390/app9091794. [CrossRef]
23. Geng, L.; Hu, Y.; Xiao, Z.; Xi, J. Fertility Detection of Hatching Eggs Based on a Convolutional Neural Network. *Appl. Sci.* **2019**, *9*, 1408. doi:10.3390/app9071408. [CrossRef]
24. Li, M.; Li, O.; Liu, G.; Zhang, C. An Automatic Modulation Recognition Method with Low Parameter Estimation Dependence Based on Spatial Transformer Networks. *Appl. Sci.* **2019**, *9*, 1010. doi:10.3390/app9051010. [CrossRef]
25. Zhuang, Z.; Lv, H.; Xu, J.; Huang, Z.; Qin, W. A Deep Learning Method for Bearing Fault Diagnosis through Stacked Residual Dilated Convolutions. *Appl. Sci.* **2019**, *9*, 1823. doi:10.3390/app9091823. [CrossRef]
26. Lin, W.J.; Lo, S.H.; Young, H.T.; Hung, C.L. Evaluation of Deep Learning Neural Networks for Surface Roughness Prediction Using Vibration Signal Analysis. *Appl. Sci.* **2019**, *9*, 1462. doi:10.3390/app9071462. [CrossRef]
27. Zheng, H.T.; Chen, J.Y.; Liang, N.; Sangaiah, A.K.; Jiang, Y.; Zhao, C.Z. A Deep Temporal Neural Music Recommendation Model Utilizing Music and User Metadata. *Appl. Sci.* **2019**, *9*, 703. doi:10.3390/app9040703. [CrossRef]
28. Kim, C.; Son, Y.; Youm, S. Chronic Disease Prediction Using Character-Recurrent Neural Network in the Presence of Missing Information. *Appl. Sci.* **2019**, *9*, 2170. doi:10.3390/app9102170. [CrossRef]
29. Karpathy, A. Multi-Layer Recurrent Neural Networks (lstm, gru, rnn) for Character-Level Language Models in Torch, 2015. Available online: https://github.com/billzorn/mtg-rnn (accessed on 4 April 2020).
30. Han, X.W.; Zheng, H.T.; Chen, J.Y.; Zhao, C.Z. Diverse Decoding for Abstractive Document Summarization. *Appl. Sci.* **2019**, *9*, 386. doi:10.3390/app9030386. [CrossRef]
31. Zhang, Y.; Li, D.; Wang, Y.; Fang, Y.; Xiao, W. Abstract Text Summarization with a Convolutional Seq2seq Model. *Appl. Sci.* **2019**, *9*, 1665. doi:10.3390/app9081665. [CrossRef]

32. Stefanovič, P.; Kurasova, O.; Štrimaitis, R. The N-Grams Based Text Similarity Detection Approach Using Self-Organizing Maps and Similarity Measures. *Appl. Sci.* **2019**, *9*, 1870. doi:10.3390/app9091870. [CrossRef]
33. Fentaw, H.W.; Kim, T.H. Design and Investigation of Capsule Networks for Sentence Classification. *Appl. Sci.* **2019**, *9*, 2200. doi:10.3390/app9112200. [CrossRef]
34. Hinton, G.E.; Krizhevsky, A.; Wang, S.D. Transforming auto-encoders. In *International Conference on Artificial Neural Networks*; Springer: Berlin/Heidelberg, Germany, 2011; pp. 44–51.
35. Bologna, G. A Simple Convolutional Neural Network with Rule Extraction. *Appl. Sci.* **2019**, *9*, 2411. doi:10.3390/app9122411. [CrossRef]
36. Yang, H.; Liu, Q.; Zhou, S.; Luo, Y. A Spam Filtering Method Based on Multi-Modal Fusion. *Appl. Sci.* **2019**, *9*, 1152. doi:10.3390/app9061152. [CrossRef]
37. Wei, C.; Ni, F.; Chen, X. Obtaining Human Experience for Intelligent Dredger Control: A Reinforcement Learning Approach. *Appl. Sci.* **2019**, *9*, 1769. doi:10.3390/app9091769. [CrossRef]
38. Radac, M.B.; Precup, R.E. Data-Driven Model-Free Tracking Reinforcement Learning Control with VRFT-based Adaptive Actor-Critic. *Appl. Sci.* **2019**, *9*, 1807. doi:10.3390/app9091807. [CrossRef]

© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).