Deep Learning for Anomaly Detection

Ruoying Wang, Kexin Nie, Tie Wang, Yang Yang, Bo Long
LinkedIn, Mountain View, CA
{ruowang,knie,tiewang,yyang,blong}@linkedin.com

ABSTRACT

Anomaly detection has been widely studied and used in diverse applications. Building an effective anomaly detection system requires the researchers/developers to learn the complex structure from noisy data, identify the dynamic anomaly patterns and detect anomalies while lacking sufficient labels. Recent advancement in deep learning techniques has made it possible to largely improve anomaly detection performance compared to the classical approaches. This tutorial will help the audience gain a comprehensive understanding of deep learning-based anomaly detection techniques in various application domains. First, it introduces what is the anomaly detection problem, the approaches taken before the deep model era and the challenges it faced. Then it surveys the state-of-the-art deep learning models extensively and discusses the techniques used to overcome the limitations from traditional algorithms. Second to last, it studies deep model anomaly detection techniques in real world examples from LinkedIn production systems. The tutorial concludes with a discussion of future trends.

CCS CONCEPTS

• **Computing methodologies → Anomaly detection**; **Neural networks**; • **Theory of computation → Semi-supervised learning**.

KEYWORDS

Anomaly Detection; Deep Learning

ACM Reference Format:
Ruoying Wang, Kexin Nie, Tie Wang, Yang Yang, Bo Long. 2020. Deep Learning for Anomaly Detection. In *The Thirteenth ACM International Conference on Web Search and Data Mining (WSDM ’20), February 3–7, 2020, Houston, TX, USA.* ACM, New York, NY, USA, 3 pages. https://doi.org/10.1145/3336191.3371876

1 INTRODUCTION

Anomaly detection is important in various applications ranging from intrusion detection [14, 27], fraud detection [8, 39–41], to medical diagnosis [12, 13, 18, 21, 32] and large scale sensor data from the Internet of Things [5, 17, 26]. The goal of anomaly detection is to identify rare abnormal data patterns that deviate from the majority of the data. The anomaly patterns are difficult to detect, due to high dimensional data structure (e.g. image and text) and temporal pattern over time. In addition, several new applications require detecting anomalies from large scale of data. It becomes increasingly challenging to apply traditional models, which often fail to identify anomalies in these cases. As we will show in this tutorial, deep learning models have successfully improved the performance of anomaly detection in the face of these challenges.

In this tutorial, we summarize the cutting-edge deep learning techniques used in various applications to detect anomalies. We first introduce anomaly detection task, and then give an overview of the traditional techniques used to detect anomalies such as statistical models, clustering, and one-class classification. We will talk about the challenges and the opportunities for more advanced algorithms.

Then we focus on introducing the state-of-the-art deep anomaly detection algorithms. In deep model anomaly detection techniques, we cover two fundamental tasks: 1) learning normal representations from complex data, where RNN, LSTM, Auto-Encoder [22, 44], GAN [9, 10, 20, 23, 33, 36] and their variations [11, 16, 24, 25, 37, 42] are widely adopted for sequential data such as text, audio [30] and time series. CNN plays a major role for non sequential data such as images [31], network and sensors; 2) detecting anomalies, while we summarize the techniques used to effectively detect anomalies based on reconstruction errors, reconstruction probabilities [3, 35, 38] and using one class NN [6]. Semi-supervised learning techniques [1, 2, 9, 10, 15, 19, 20, 23–25, 33, 36, 37] and transfer learning [4, 15] are presented, which are used to compensate for sparse anomaly labels. In deep anomaly detection architectures, we introduce the architecture of deep learning anomaly detection model including hybrid models [28, 29, 34] and spatial temporal network [7, 43].

Second to last, we evaluate deep learning methodologies on several publicly available data sets. What’s more, we illustrate the end-to-end anomaly detection product at LinkedIn, by sharing our experiences for multivariate time series deep anomaly detection, multi-step horizon forecasting and pattern-based deep anomaly detection. In the end, we highlight several important future trends.

Targeted Audience This tutorial is suitable for academic and industrial researchers, graduate students, and practitioners. After the tutorial, we expect the audience to have learnt the key concepts and principles of applying the state-of-the-art deep learning models for anomaly detection, and gained real-world experiences through illustrative examples.

WSDM ’20, February 3–7, 2020, Houston, TX, USA

ACM ISBN 978-1-4503-6822-3/20/02.
https://doi.org/10.1145/3336191.3371876

2 TUTORIAL OUTLINE

2.1 Introduction (30 min)

(1) Overview of Anomaly Detection
(2) Traditional Techniques
(3) Anomaly Detection Application and Challenges

2.2 Deep Learning for Anomaly Detection (90 min)

(1) Deep Model Preliminaries
(2) Deep Model Anomaly Detection Techniques
  (a) Learning Complex Normal Data Representations
    (i) Sequential Data
    (ii) Non-Sequential Data
  (b) Detecting Anomalies
(3) Compensating for Sparse Anomaly Labels
  (a) Semi-Supervised Learning
  (b) Transfer Learning
(4) Deep Anomaly Detection Architecture
  (a) Hybrid Model
  (b) Spatial Temporal Network

2.3 Real-world Example Using Deep Anomaly Detection (50 min)

(1) Performance Evaluation on Public Data Sets
(2) Deep Model for Anomaly Detection at LinkedIn
  (a) Anomaly Detection of Multivariate Time Series
  (b) Multi-step Horizon Forecast
  (c) Pattern-based Anomaly Detection

2.4 Conclusions and Future Trends (10 min)

3 PRESENTERS' BIOGRAPHY

Dr. Ruoying Wang is an AI software engineer at LinkedIn. She works on applying deep learning and statistical models for anomaly detection and capacity planning. She obtained her Ph.D. in Economics degree from UBC, focusing on empirical causal analysis with applications in International Trade. She is excited to develop deep learning algorithms for anomaly detection in production at LinkedIn.

Kexin Nie is a Sr. AI software engineer at LinkedIn, where she leads the effort to monitor AI models' online performance drift and automatic diagnose issues' root causes at scale. She has been working in the field of anomaly detection for 2+ years and launched several algorithms in LinkedIn's health monitoring service (ThirdEye). Before this, she worked for IBM to optimize its E-commercial ads' tagging. She obtained her Master of Statistics from Stanford University.

Dr. Tie Wang leads the AI Quality Foundation team at LinkedIn. His team owns the anomaly detection algorithm library, that supports anomaly detection for over 20 LinkedIn products. He has broad interests in machine learning/AI and its applications. He has 12 years of R&D experience at Apple, Microsoft, LinkedIn on anomaly detection, query understanding, commercial and web search ranking algorithms and systems. He received Ph.D. in Computer Science from Arizona State University. He has published in top journals and conferences including KDD, IEEE Transaction on Signal Process.

Dr. Yang Yang is a Senior Staff Software Engineer and Tech Lead at LinkedIn. Before joining LinkedIn, Yang worked at Yahoo! Labs as a Scientist. She obtained her Ph.D. degree at Department of Statistics, University of Michigan. She has produced various papers and patents on applying statistical methods and machine learning approaches to real data problem involving large scale data. She has published in conferences and journals including KDD, WWW, PAM, Statistical Analysis and Data Mining, The Canadian Journal of Statistics, IIE Transactions on Healthcare Systems Engineering, and Statistical Analysis for High-Dimensional Data.

Dr. Bo Long is a Director of AI Engineering at LinkedIn, leading LinkedIn's AI Foundations team. He has 15 years of experience in data mining and machine learning with applications to web search, recommendation, and social network analysis. He holds dozens of innovations and has published peer reviewed papers in top conferences and journals including ICML, KDD, ICDM, AAAI, SDM, CIKM, and KAIS. He has served as reviewers, workshops co-organizers, conference organizer committee members, and area chairs for multiple conferences, including KDD, NIPS, SIGIR, ICML, SDM, CIKM, JSM etc.

REFERENCES

[1] Samet Akcay, Amir Atapour-Abarghouei, and Toby P Breckon. 2018. Ganomaly: Semi-supervised anomaly detection via adversarial training. In *Asian Conference on Computer Vision*. Springer, 622–637.
[2] Samet Akçay, Amir Atapour-Abarghouei, and Toby P Breckon. 2019. Skip-ganomaly: Skip connected and adversarially trained encoder-decoder anomaly detection. *arXiv preprint arXiv:1901.08954* (2019).
[3] Jinwon An and Sungzoon Cho. 2015. Variational autoencoder based anomaly detection using reconstruction probability. *Special Lecture on IE* 2, 1 (2015).
[4] Jerone TA Andrews, Thomas Tanay, Edward J Morton, and Lewis D Griffin. 2016. Transfer representation-learning for anomaly detection. In *Proc. ICML*. 1–5.
[5] Raghavendra Chalapathy and Sanjay Chawla. 2019. Deep Learning for Anomaly Detection: A Survey. *CoRR* abs/1901.03407 (2019). arXiv:1901.03407 http://arxiv.org/abs/1901.03407
[6] Raghavendra Chalapathy, Aditya Krishna Menon, and Sanjay Chawla. 2018. Anomaly detection using one-class neural networks. *arXiv preprint arXiv:1802.06360* (2018).
[7] Dan Chianucci and Andreas Savakis. 2016. Unsupervised change detection using spatial transformer networks. In *2016 IEEE Western New York Image and Signal Processing Workshop (WNYISPW)*. IEEE, 1–5.
[8] Alae Chouiekh and EL Hassane Ibn EL Haj. 2018. Convnets for fraud detection analysis. *Procedia Computer Science* 127 (2018), 133–138.
[9] Antonia Creswell and Anil Anthony Bharath. 2018. Inverting the generator of a generative adversarial network. *IEEE transactions on neural networks and learning systems* (2018).
[10] Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. 2014. Generative adversarial nets. In *Advances in neural information processing systems*. 2672–2680.
[11] Yifan Guo, Weixian Liao, Qianlong Wang, Lixing Yu, Tianxi Ji, and Pan Li. 2018. Multidimensional time series anomaly detection: A gru-based gaussian mixture variational autoencoder approach. In *Asian Conference on Machine Learning*. 97–112.
[12] Dimitris K Iakovidis, Spiros V Georgakopoulos, Michael Vasilakakis, Anastasios Koulaouzidis, and Vassilis P Plagianakos. 2018. Detecting and locating gastrointestinal anomalies using deep learning and iterative cluster unification. *IEEE transactions on medical imaging* 37, 10 (2018), 2196–2210.
[13] Abhyuday N Jagannatha and Hong Yu. 2016. Bidirectional RNN for medical event detection in electronic health records. In *Proceedings of the conference. Association for Computational Linguistics. North American Chapter. Meeting*, Vol. 2016. NIH Public Access, 473.
[14] Ahmad Javaid, Quamar Niyaz, Weiqing Sun, and Mansoor Alam. 2016. A deep learning approach for network intrusion detection system. In *Proceedings of the 9th EAI International Conference on Bio-inspired Information and Communications Technologies (formerly BIONETICS)*. ICST (Institute for Computer Sciences, Social-Informatics and âĂę, 21–26.
[15] Durk P Kingma, Shakir Mohamed, Danilo Jimenez Rezende, and Max Welling. 2014. Semi-supervised learning with deep generative models. In *Advances in neural information processing systems*. 3581–3589.
[16] Diederik P Kingma and Max Welling. 2014. Stochastic gradient VB and the variational auto-encoder. In *Second International Conference on Learning Representations, ICLR*.
[17] B Kiran, Dilip Thomas, and Ranjith Parakkal. 2018. An overview of deep learning based methods for unsupervised and semi-supervised anomaly detection in videos. *Journal of Imaging* 4, 2 (2018), 36.

[18] Daniel Lasaga and Prakash Santhana. 2018. Deep learning to detect medical treatment fraud. In *KDD 2017 Workshop on Anomaly Detection in Finance*. 114–120.

[19] Chunyuan Li, Hao Liu, Changyou Chen, Yuchen Pu, Liqun Chen, Ricardo Henao, and Lawrence Carin. 2017. Alice: Towards understanding adversarial learning for joint distribution matching. In *Advances in Neural Information Processing Systems*. 5495–5503.

[20] Dan Li, Dacheng Chen, Lei Shi, Baihong Jin, Jonathan Goh, and See-Kiong Ng. 2019. MAD-GAN: Multivariate anomaly detection for time series data with generative adversarial networks. *arXiv preprint arXiv:1901.04997* (2019).

[21] Geert Litjens, Thijs Kooi, Babak Ehteshami Bejnordi, Arnaud Arindra Adiyoso Setio, Francesco Ciompi, Mohsen Ghafoorian, Jeroen Awm Van Der Laak, Bram Van Ginneken, and Clara I Sánchez. 2017. A survey on deep learning in medical image analysis. *Medical image analysis* 42 (2017), 60–88.

[22] Pankaj Malhotra, Anusha Ramakrishnan, Gaurangi Anand, Lovekesh Vig, Puneet Agarwal, and Gautam Shroff. 2016. LSTM-based encoder-decoder for multi-sensor anomaly detection. *arXiv preprint arXiv:1607.00148* (2016).

[23] Federico Di Mattia, Paolo Galeone, Michele De Simoni, and Emanuele Ghelfi. 2019. A Survey on GANs for Anomaly Detection. *CoRR* abs/1906.11632 (2019). arXiv:1906.11632 http://arxiv.org/abs/1906.11632

[24] Mehdi Mirza and Simon Osindero. 2014. Conditional generative adversarial nets. *arXiv preprint arXiv:1411.1784* (2014).

[25] Takeru Miyato, Toshiki Kataoka, Masanori Koyama, and Yuichi Yoshida. 2018. Spectral normalization for generative adversarial networks. *arXiv preprint arXiv:1802.05957* (2018).

[26] Mehdi Mohammadi, Ala Al-Fuqaha, Sameh Sorour, and Mohsen Guizani. 2018. Deep learning for IoT big data and streaming analytics: A survey. *IEEE Communications Surveys & Tutorials* 20, 4 (2018), 2923–2960.

[27] Mutahir Nadeem, Ochaun Marshall, Sarbjit Singh, Xing Fang, and Xiaohong Yuan. 2016. Semi-supervised deep neural network for network intrusion detection. (2016).

[28] Trong Nguyen Nguyen and Jean Meunier. 2019. Hybrid Deep Network for Anomaly Detection. arXiv:cs.CV/1908.06347

[29] Miguel Nicolau, James McDermott, et al. 2016. A hybrid autoencoder and density estimation model for anomaly detection. In *International Conference on Parallel Problem Solving from Nature*. Springer, 717–726.

[30] Timothy J O'Shea, T Charles Clancy, and Robert W McGwier. 2016. Recurrent neural radio anomaly detection. *arXiv preprint arXiv:1611.00301* (2016).

[31] Mohammad Sabokrou, Mohsen Fayyaz, Mahmood Fathy, Zahra Moayed, and Reinhard Klette. 2018. Deep-anomaly: Fully convolutional neural network for fast anomaly detection in crowded scenes. *Computer Vision and Image Understanding* 172 (2018), 88–97.

[32] Daisuke Sato, Shouhei Hanaoka, Yukihiro Nomura, Tomomi Takenaga, Soichiro Miki, Takeharu Yoshikawa, Naoto Hayashi, and Osamu Abe. 2018. A primitive study on unsupervised anomaly detection with an autoencoder in emergency head CT volumes. In *Medical Imaging 2018: Computer-Aided Diagnosis*, Vol. 10575. International Society for Optics and Photonics, 105751P.

[33] Thomas Schlegl, Philipp Seeböck, Sebastian M Waldstein, Ursula Schmidt-Erfurth, and Georg Langs. 2017. Unsupervised anomaly detection with generative adversarial networks to guide marker discovery. In *International Conference on Information Processing in Medical Imaging*. Springer, 146–157.

[34] Hongchao Song, Zhuqing Jiang, Aidong Men, and Bo Yang. 2017. A hybrid semi-supervised anomaly detection model for high-dimensional data. *Computational intelligence and neuroscience* 2017 (2017).

[35] Haowen Xu, Wenxiao Chen, Nengwen Zhao, Zeyan Li, Jiahao Bu, Zhihan Li, Ying Liu, Youjian Zhao, Dan Pei, Yang Feng, et al. 2018. Unsupervised anomaly detection via variational auto-encoder for seasonal kpis in web applications. In *Proceedings of the 2018 World Wide Web Conference*. International World Wide Web Conferences Steering Committee, 187–196.

[36] Houssam Zenati, Chuan Sheng Foo, Bruno Lecouat, Gaurav Manek, and Vijay Ramaseshan Chandrasekhar. 2018. Efficient gan-based anomaly detection. *arXiv preprint arXiv:1802.06222* (2018).

[37] Houssam Zenati, Manon Romain, Chuan Sheng Foo, Bruno Lecouat, and Vijay Ramaseshan Chandrasekhar. 2018. Adversarially Learned Anomaly Detection. *CoRR* abs/1812.02288 (2018). arXiv:1812.02288 http://arxiv.org/abs/1812.02288

[38] Shuangfei Zhai, Yu Cheng, Weining Lu, and Zhongfei Zhang. 2016. Deep structured energy based models for anomaly detection. *arXiv preprint arXiv:1605.07717* (2016).

[39] Zhaohui Zhang, Xinxin Zhou, Xiaobo Zhang, Lizhi Wang, and Pengwei Wang. 2018. A model based on convolutional neural network for online transaction fraud detection. *Security and Communication Networks* 2018 (2018).

[40] Panpan Zheng, Shuhan Yuan, Xintao Wu, Jun Li, and Aidong Lu. 2019. One-class adversarial nets for fraud detection. In *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 33. 1286–1293.

[41] Yu-Jun Zheng, Xiao-Han Zhou, Wei-Guo Sheng, Yu Xue, and Sheng-Yong Chen. 2018. Generative adversarial network based telecom fraud detection at the receiving bank. *Neural Networks* 102 (2018), 78–86.

[42] Chong Zhou and Randy C Paffenroth. 2017. Anomaly detection with robust deep autoencoders. In *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. ACM, 665–674.

[43] Shifu Zhou, Wei Shen, Dan Zeng, Mei Fang, Yuanwang Wei, and Zhijiang Zhang. 2016. Spatial–temporal convolutional neural networks for anomaly detection and localization in crowded scenes. *Signal Processing: Image Communication* 47 (2016), 358–368.

[44] Bo Zong, Qi Song, Martin Renqiang Min, Wei Cheng, Cristian Lumezanu, Daeki Cho, and Haifeng Chen. 2018. Deep Autoencoding Gaussian Mixture Model for Unsupervised Anomaly Detection. (2018). https://openreview.net/forum?id=BJJLHbb0-