Use of Generative Adversarial Networks (GANs) in Anomaly Detection

Cindy Wang

Abstract

In a world growing increasingly digital, the applications of anomaly detection are expanding, particularly in cybersecurity and in real-world monitoring. The development of anomaly detection models comes with various challenges. In current literature, a plethora of solutions are explored, including the use of generative adversarial networks in expanding the scope of anomaly detection. However, the use of GANs in this field has not been fully explored. This paper aims to review and analyze current solutions for improving the development of detection algorithms, along with proposing future areas for research and improvements to be made.

Introduction

Anomaly Detection

Anomaly detection is a widely used technique that aims at identifying outliers from normal behavior in datasets [1]. It has wide applications in fields like cybersecurity, where malicious attacks can be identified, and banking, where fraudulent activity can be pinpointed. However, current challenges associated with the field of anomaly detection are the large amount of resources it takes to label data and the rapidly changing "normal" values in real-world multivariate data that make it difficult to identify inconsistent behavior [2]. These problems make it difficult for businesses and researchers to effectively scale their anomaly detection algorithms and must be tackled to improve on current detection methods.

Generative Adversarial Networks

One possible solution to these problems in anomaly detection is the use of generative adversarial networks or GANs. GANs can learn the patterns of input data and generate new datasets based on inputs. Because GANs can learn unsupervised with a high degree of accuracy, unlabeled datasets can be used in generating new information, which is highly beneficial for anomaly detection uses [3].

Nevertheless, there are limitations to the extent and accuracy of GANs, given that the data being generated is not taken from real sources. Implicit biases also affect results from GANs [4].

Current Research

MAD-GAN

There are various current solutions being explored to improve on parts of anomaly detection algorithms using GANs. The first of these is MAD-GAN. Unlike traditional anomaly detection algorithms that considers only univariate streams of data, and then combines the findings for application to multivariate datasets, MAD-GAN takes into consideration all variables in the dataset [5]. This proposed solution was shown to be effective in real-world cybersecurity applications.

One aspect that this article does not focus on is the use of GANs for new data generation, which highlights an area for improvement in anomaly detection.

Marker Discovery

Becuase of the labeling of data that is often needed in anomaly detection, Schlegl et al. developed a methodology for generating realistic-looking images of the retina, along with creating labels for this data for highly-accurate anomaly detection [6]. The algorithm was shown to have a high degree of accuracy in detecting retinal anomalies, such as retinal fluid or hyperreflective foci.

Proposed Solution

RenderGAN: Creating Labeled Datasets

Because anomaly detection requires large amounts of data for optimal accuracy and in the testing of algorithms, GANs are an optimal solution for generating labeled datasets. Currently, state-of-the-art generation models like RenderGAN are able to generate and label realistic image datasets with an accuracy of 96% [7]. Because this model generates realistic background, lighting, and details from unlabeled data, this work has large applications within real-world image-based detection algorithms, from which models are often trained without labeled data [8]. Possible fields of application for RenderGAN include surveillance and crime detection, manufacturing quality control, and identification of agricultural anomalies.

GANs for Beginning Training

Because anomaly detection can be a crucial component in companies, such as cybersecurity or banking startups, it is often necessary to begin anomaly detection earlier. However, without a large amount of normalized data, it is very difficult to determine which data points are anomalous. In that aspect, GANs can likely be used to generate normal data as systems and behavior evolve over time.

Limitations

One of the main limitations of GANs that may affect its usage in anomaly detection is mode collapse. Many GANs tend to be prone to mode collapse, meaning that they do not truly cover the entire scope of data and miss edge cases [9]. While solutions, such as VEEGAN, developed by Srivastava et al. specifically to reduce mode collapse, have been created, there is still a high chance of missing many anomaly points and edge cases.

A proposed solution to this is to develop and train a GAN model with a specific focus on edge cases, to ensure that anomalous data points are covered.

References

[1] Chandola V., Banerjee A., Kumar V., “Anomaly detection: A survey”, ACM Computing Surveys, Volume 41, Issue 3, July 2009. https://dl.acm.org/doi/10.1145/1541880.1541882

[2] Michailidis G., “Challenges for Anomaly Detection in Large-Scale Cyber-Physical Systems”, Harvard Data Science Review, Volume 5, Issue 1, March 2023. https://hdsr.mitpress.mit.edu/pub/c5ikncc8/release/1

[3] Brownlee J., “A Gentle Introduction to Generative Adversarial Networks (GANs)”, Machine Learning Mastery, July 2019. https://machinelearningmastery.com/what-are-generative-adversarial-networks-gans/

[4] Jain N., et al., “Imperfect ImaGANation: Implications of GANs exacerbating biases on facial data augmentation and snapchat face lenses”, Artificial Intelligence, Volume 304, March 2022. https://www.sciencedirect.com/science/article/abs/pii/S0004370221002034

[5] Li D., et al, “MAD-GAN: Multivariate Anomaly Detection for Time Series Data with Generative Adversarial Networks”, International Conference on Artificial Neural Networks, September 2019. https://link.springer.com/chapter/10.1007/978-3-030-30490-4_56

[6] Schlegl T., “Unsupervised Anomaly Detection with Generative Adversarial Networks to Guide Marker Discovery”, International Conference on Information Processing in Medical Imaging, May 2017. https://link.springer.com/chapter/10.1007/978-3-319-59050-9_12

[7] Sixt L., Wild B., Landgraf T., “RenderGAN: Generating Realistic Labeled Data”, Frontiers in Robotics and AI, Volume 5, June 2018. https://www.frontiersin.org/articles/10.3389/frobt.2018.00066/full

[8] Gornitz N., et al., “Towards Supervised Anomaly Detection”, Journal of Artificial Intelligence Research, Volume 46, February 2013. https://arxiv.org/pdf/1401.6424.pdf

[9] Srivastava A., et al., “VEEGAN: Reducing Mode Collapse in GANs using Implicit Variational Learning”, 31st Conference on Neural Information Processing Systems, 2017. https://proceedings.neurips.cc/paper_files/paper/2017/file/44a2e0804995faf8d2e3b084a1e2db1d-Paper.pdf