Deep Learning for Anomaly Detection: A Survey

1 Abstract

https://arxiv.org/abs/1901.03407

We examine different deep anomaly detection (DAD) techniques, group them, and look at applications.

2 Introduction

An anomaly is a point that deviates so significantly from other observations as to arouse suspicion that it was generated by a different mechanism (Hawkins 1980). When you have a lot of data, deep learning methods can beat traditional approaches.

3 What are Anomalies?

Anomalies are points that are located far away from the regular data distribution.

4 What are Novelties?

These are points that are not wildly different from existing data points, but have some unique aspect of them that is interesting (e.g. a white tiger is an anomaly in a dataset of animals).

5 Motivation and Challenges: Deep anomaly detection (DAD) techniques

Traditional methods are hard to scale and don't work well for image and sequence (e.g. text) data. It's also a pain trying to engineer good features.

6 Related Work

There are other survey papers that look at DAD for fraud detection, medicine, and more.

7 Our Contributions

This paper considers DAD comprehensively, rather than focusing on a specific domain. It also looks at hybrid and one-class models.

8 Organization

Nothing useful here.

9 Different aspects of deep learning-based anomaly detection

What is the nature of the input data? Is it text, image, voice, video, etc?

Do you have any labels? Are you doing supervised, semi-supervised, or unsupervised learning? Supervision is hard because we usually don't have labels, and, even if we do, dealing with the class imbalance is tricky. Let's ignore supervised DAD for this paper. In the semi-supervised case, we have normal instances only. We can train an autoencoder on this data and consider outliers to be points with high reconstruction error. The unsupervised case is the most common one, and DAD beats SVM, PCA, and isolation forests. Autoencoders are our main workhorse here, but belief nets and LSTMs can also help.

Deep Hybrid Models (DHM) train an autoencoder and then use it as a feature extractor that gets fed into a traditional anomaly detection method. One Class Neural Networks (OC-NN) have a special loss function designed for anomaly detection and they learn a good representation for normal data.

There are three kinds of anomalies. Point anomalies are points that are very different than other points. A contextual anomaly is a point that is different than other points that are near (in space and/or time) it. Collective anomalies (or group anomalies) are groups of points that are different from the rest of the dataset.

Anomaly detection methods either predict an anomaly score or binary label for each point.

10 Applications of Deep Anomaly Detection

Intrusion detection comes in two flavors - host intrusion detection or network intrusion detection. Some methods use signatures of known attacks, but this cannot find new attacks. So, we need anomaly detection systems.

Fraud detection is another application. For credit card fraud, it's typical to keep a per-user profile and compare new transactions against it, but this does not scale, so DAD methods are useful. Mobile cellular network fraud, insurance fraud, and healthcare fraud are tough to detect because you need to design good features, so DAD helps here too.

Malware detection usually involves clustering and feature extraction, but DAD approaches can help out here. Medical anomaly detection is a good use of DAD. It can even detect anomalies in social networks. Detecting anomalies in log data also works well with DAD because neural nets can learn well from text data.

DAD can also detect anomalies in Internet of Things systems and industrial systems like turbines.

DAD can also be used for univariate and multivariate time series. We need to take care to handle nonstationary data distributions.

Finally, we can use DAD for detecting anomalies in video surveillance data.

11 Deep Anomaly Detection (DAD) Models

Supervised DAD works better than semi-supervised or unsupervised methods if you have a large enough labeled dataset.

Semi-supervised techniques learn a boundary around the normal class. Autoencoders and Generative Adversarial Networks (GANs) work well here.

Hybrid approaches train a feature extractor network (an autoencoder) and then feed the result to a traditional outlier detection algorithm. They key problem here is that the model is not end-to-end differentiable. You can overcome this limitation with one-class neural networks that have a special loss function for outlier detection (NOTE: I think the author of this survey is championing this approach partly because they also wrote a paper on one-class neural networks).

For unsupervised anomaly detection, autoencoders are the primary workhorse. Be careful in picking the dimension of the compressed representation of the autoencoder. Also, be aware that unsupervised methods do not work well on noisy datasets.

Recent research has explored transfer learning for DAD, zero-shot learning (i.e. recognizing new classes given some metadata about them), ensembling DAD (by randomly varying the autoencoder architecture), clustering (i.e. running a word2vec-like algorithm and clustering the results), and deep reinforcement learning DAD (very new).

12 Deep neural network architectures for locating anomalies

Deep Belief Nets and Restricted Boltzmann Machines build up their architecture one layer at a time in an unsupervised way. They can be a good way to generate feature extractors.

Spatiotemporal networks combine convolutional neural networks (CNNs) and Long Short Term Memory (LSTMs) to extract spatiotemporal features.

Sum-product networks are directed acyclic graphs with variables at the leaves and internal nodes (and weighted edges) to represent sums and products.

Word2Vec is a good way to learn embeddings vectors for data points. They are useful in situation where we have co-occurence information about data points (e.g. they are near each other in space or time).

Generative models like Variational Autoencoders (VAEs) and Generative Adversarial Networks (GANs) seem promising. Adversarial Autoencoders also seem promising, but traditional methods like k-nearest neighbors sometimes beat them.

CNNs have been very useful in other domains, but people are still researching them for outlier detection. Sequence models, like the LSTM, have also been studied.

Autoencoders do a good job at learning the normal class and produce high reconstruction errors for outliers. They are combined with CNNs, or LSTMs, when you are processing image or sequence data, respectively. They struggle on high dimensional datasets though.

13 Relative Strengths and Weakness: Deep Anomaly Detection Methods

If you have lots of data, use a supervised model. Otherwise use a semi-supervised or unsupervised approach. One-class Neural Networks beat hybrid models.

14 Conclusion

We looked at various anomaly detection applications and deep learning models for them.