Utilizing Deviation Networks for Anomaly Detection

V.Mounika[1],

[1]Asst Professor, Department of CSE, Koneru Lakshmaiah Education Foundation,
Vaddeswaram, AP, India.. vmounika@kluniversity.in

Dr.N.Raghavendra Sai[2]

[1]Asst Professor, Department of CSE, Koneru Lakshmaiah Education Foundation,
Vaddeswaram, AP, India.. nallagatlaraghavendra@kluniversity.in

Abstract:

Anomaly detection acts significant role in various domains, including cyber security, fraud detection, and industrial monitoring. Traditional approaches depend on handcrafted features and assumptions about data distributions, limiting their effectiveness and adaptability. This paper introduces Deviation Networks, a deep learning-based technique useful for abnormality identification that leverages the power of deep neural networks to learn complex representations directly from data. Deviation Networks employ an encoder-decoder architecture combined with deviation-based loss functions to capture normal patterns in the training data and identify deviations indicative of anomalies. This paper provides an in-depth exploration of Deviation Networks, including their architecture, training procedure, and evaluation metrics. Furthermore, we present experimental results on benchmark datasets, demonstrating the effectiveness and superiority of Deviation Networks compared to traditional anomaly detection methods. We also discuss practical considerations, such as data pre-processing, hyper parameter tuning, and deployment strategies. Overall, this paper showcases the potential of identification of abnormalities using Dense Networks and highlights their significance in addressing real-world anomaly detection challenges.

Keywords: Anomaly detection, deep learning, Deviation Networks, encoder-decoder architecture, deviation-based loss functions, cyber security, fraud detection, industrial monitoring, benchmark datasets, data pre-processing, hyper parameter tuning, deployment strategies.

1. Introduction

Anomaly detection, also known as outlier detection, is a critical task in various domains, including cybersecurity, fraud detection, industrial monitoring, and healthcare. The goal of anomaly detection is to identify patterns or instances that significantly deviate from the expected or normal behavior within a dataset. Traditional methods for anomaly detection often rely on handcrafted features and make assumptions about the underlying data distribution, in capturing complex anomalies.

In the modern world deep learning has emerged as a powerful technique for learning representations directly from data, enabling the making of more robust and adaptive anomaly detection methods. Deep learning models, such as autoencoders and generative adversarial networks (GANs), have demonstrated promising results in capturing intricate patterns and anomalies in various types of data.

The motto of this study involves Deviation Networks, a novel approach for deep anomaly detection that leverages the capabilities of deep neural networks to learn complex representations and identify anomalies. Deviation Networks employ an encoder-decoder architecture that learns to reconstruct the input data, in align of capturing the common patterns within the dataset. The models are trained using deviation-based loss functions, which focus on quantifying the deviation between the reconstructed output and the original input. By optimizing these loss functions, the Deviation Networks can effectively identify anomalies as instances with significant reconstruction errors.

This paper contributes are as follows:

a. An in-depth exploration of Deviation Networks, including their architecture, training procedure, and deviation-based loss functions.
b. Evaluation of Deviation Networks on benchmark datasets and comparison with traditional anomaly detection methods, showcasing their superior performance.
c. Discussion of practical considerations such as data preprocessing, hyperparameter tuning, and deployment strategies to guide the implementation of Deviation Networks in real-world applications.
d. Illustration of the potential applications of Deviation Networks in cybersecurity, fraud detection, industrial monitoring, and other relevant domains.
e. Identification of challenges and future directions in the field of deep anomaly detection, including scalability, interpretability, and handling concept drift.

2. Anomaly Detection Methods

Anomaly detection has been a focal point of research for numerous years, and various methods have been developed to tackle this task. Traditional anomaly detection methods can be categorized into statistical, distance-based, clustering, and rule-based approaches. While these methods have been effective in certain scenarios, they often rely on assumptions about the data distribution and may struggle to capture complex anomalies or adapt to changing data patterns[1].

Fig 1: **Anomaly Detection Methods**

2.1 Statistical Methods:

Statistical methodologies that are used for abnormality detection often assume that the data follows a specific probability distribution, such as Gaussian or exponential. These methods utilize statistical measures like mean, variance, and probability density functions to identify instances that deviate significantly from the expected distribution. Examples of statistical methods include Gaussian Mixture Models (GMM), z-score, and percentile-based approaches. However, these methods may fail to capture anomalies that do not conform to the assumed distribution or when the data is high-dimensional and complex.

2.2 Distance-Based Methods:

Distance-based methods identify anomalies based on the notion of distance or dissimilarity between data points. These methods calculate distances, such as Euclidean distance or Mahalanobis distance, and identify instances that are significantly distant from the rest of the data points. One common distance-based approach is the k-nearest neighbors (k-NN) algorithm, where anomalies are considered as instances with a low number of nearby neighbors. Distance-based methods can be effective for detecting isolated anomalies but may struggle with high-dimensional data or when the data distribution is complex and nonlinear.

2.3 Clustering Methods:

Clustering-based anomaly detection approaches aim to identify instances that do not belong to any specific cluster. These methods partition the data into clusters and consider instances that do not fit well within any cluster as anomalies. One popular algorithm in clustering for anomaly detection is the k-means algorithm. However, clustering methods may suffer from the assumption that anomalies form separate clusters, which may not hold true in all cases, especially when anomalies are subtle or exhibit similar patterns to normal instances.

2.4 Rule-Based Methods:

Rule-based methods define a set of rules or thresholds based on domain knowledge or expert-defined heuristics to identify anomalies. These methods often rely on specific conditions or rules to flag instances as anomalies. For example, in network intrusion detection, rule-based methods may define specific patterns of network traffic as anomalous. Rule-based methods can be successful when there is prior familiarity with the anomalies is available, but they may struggle to handle complex or unknown anomalies that's not predefined rules.

While traditional anomaly detection While these methods have found extensive application, they are not without their constraints. in capturing complex and evolving anomalies. Deep learning-based approaches, such as Deviation Networks, have shown promise in addressing these limitations by automatically learning representations of text and capturing intricate patterns and deviations. The subsequent segments will explore the architecture and methodology of Deviation Networks for deep anomaly detection.

3. Deviation Networks

Deviation Networks are a deep learning-based strategy for anomaly detection that leverage the power of deep neural networks to learn complex representations directly from data. These networks are designed to capture normal patterns within the training data and identify deviations indicative of anomalies.

3.1 Architecture Overview:

The Design of Deviation Networks typically consists of an encoder-decoder design. The encoder of the network learns to extract meaningful representations or features from the input data. It reduces the dimensionality of the data and captures the underlying patterns. The decoder part of the network aims to reconstruct the input data from the learned representations. By comparing the reconstructed output with the original input, deviations or reconstruction errors can be quantified and used to identify anomalies.

3.2 Encoder-Decoder Design:

The encoder component of Deviation Networks typically consists of multiple layers, such as convolutional layers, recurrent layers, or fully connected layers, depending on the nature of the input data. These layers progressively learn hierarchical representations of the data, capturing both low-level and high-level features. The decoder component mirrors the structure of the encoder, with the layers reversed. It aims to reconstruct the input data from the learned representations, emphasizing the preservation of Essential data is retained while eliminating irrelevant or noisy details, ensuring a streamlined and meaningful dataset..

3.3 Deviation-Based Loss Functions:

eviation Networks employ deviation-based loss functions to quantify the difference between the reconstructed output and the original input. These loss functions capture the notion of deviation or abnormality by assigning higher loss values to instances with larger reconstruction errors. Common loss functions used in Deviation Networks include mean squared error (MSE), binary cross-entropy, or other customized loss functions tailored to the specific anomaly detection task.

3.4 Training Procedure:

The training procedure for Deviation Networks involves feeding the training data to the network and optimizing the deviation-based loss function. The network learns to minimize the reconstruction error on the training data, enabling it to capture the normal patterns present in the data. During training, both normal and anomalous instances can be used, with a focus on ensuring that the network can effectively distinguish between the two. Various optimization techniques, such as stochastic gradient descent (SGD) or Adam, can be employed to update the network parameters and fine-tune the model.

Deviation Networks can be trained in an unsupervised manner, where only normal instances are available during training, or in a semi-supervised setting, where a small portion of labeled anomalous instances are included to guide the learning process.

By Analyzing the power of deep learning and optimizing the deviation-based loss functions, Deviation Networks can effectively identify anomalies based on significant reconstruction errors[2]. In the next section, we will discuss the evaluation metrics used to assess the performance of anomaly detection methods, including Deviation Networks.

4.Evaluation Metrics

Evaluation metrics play a pivotal role in gauging the effectiveness of anomaly detection methods, including Deviation Networks. These metrics furnish quantitative evaluations of the model's ability to accurately detect anomalies and distinguish them from normal instances. Here are some commonly used evaluation metrics for anomaly detection:

4.1 Accuracy, Precision, and Recall:

Accuracy measures the overall correctness of anomaly detection by computing the ratio of accurately identified anomalies to the total number of instances. Nevertheless, relying solely on accuracy may not be sufficient for imbalanced datasets, where anomalies are rare compared to normal instances. In such cases, precision and recall are more informative metrics. Precision represents the proportion of correctly detected anomalies out of all instances flagged as anomalies, while recall calculates the proportion of correctly detected anomalies out of all actual anomalies. A balance between precision and recall is crucial, and the trade-off can be assessed using the F1 score, which combines both precision and recall into a single metric.

4.2 Receiver Operating Characteristic (ROC) Curve:

The ROC curve is a graphical representation of the true positive rate (TPR) against the false positive rate (FPR) at various threshold settings. The TPR represents the proportion of true anomalies correctly identified, while the FPR measures the proportion of normal instances incorrectly flagged as anomalies. The ROC curve provides a visual depiction of the model's performance across different threshold settings and can help determine an appropriate threshold based on the desired trade-off between TPR and FPR. The area under the ROC curve (AUC) is a widely used metric to quantify the overall performance of the model, with a higher AUC indicating better performance

4.3 Area Under the Precision-Recall Curve (AUPRC):

The precision-recall curve is another graphical representation that illustrates the trade-off between precision and recall at different threshold settings. Similar to the ROC curve, the AUPRC provides a quantitative measure of the model's performance by calculating the area under the precision-recall curve. A higher AUPRC indicates better performance in capturing anomalies while minimizing false detections.

4.4 F1 Score:

The F1 score is a harmonic mean of precision and recall and provides a balanced measure of the model's performance. It is particularly useful when dealing with imbalanced datasets where anomalies are rare. The F1 score combines precision and recall into a single metric, where a higher F1 score indicates better overall performance.

These evaluation metrics help quantify the performance of Deviation Networks and other anomaly detection methods. The choice of evaluation metrics depends on the specific requirements of the application and the characteristics of the dataset. In the next section, we will present experimental results on benchmark datasets, comparing Deviation Networks with traditional methods to showcase their effectiveness in anomaly detection.

5.Experimental Results

To evaluate the performance of Deviation Networks in anomaly detection, experiments were conducted on benchmark datasets and compared with traditional anomaly detection methods[3]. The results demonstrate the effectiveness and superiority of Deviation Networks in capturing anomalies and distinguishing them from normal instances.

5.1 Benchmark Datasets:

Several benchmark datasets were selected to evaluate the performance of Deviation Networks across different domains. These datasets encompass various types of data, such as time series data, image data, and tabular data. Examples of benchmark datasets used for evaluation include:

KDD Cup 99: A network intrusion detection dataset containing network traffic data with different types of attacks and normal traffic.

MNIST: A widely used dataset of hand-written digit images.

CIFAR-10: A dataset of 50,000 images belonging to 10 different classes, commonly used for image classification tasks.

Credit Card Fraud: A dataset containing credit card transactions, with a highly imbalanced distribution between normal and fraudulent transactions.

Numenta Anomaly Benchmark (NAB): A collection of real-world time series datasets with labeled anomalies, including temperature, power, and CPU usage data.

5.2 Comparison with Traditional Methods:

Deviation Networks were compared with traditional anomaly detection methods, including statistical approaches, distance-based methods, clustering methods, and rule-based methods.

The evaluation focused on metrics such as accuracy, precision, recall, F1 score, ROC curve, AUC, and AUPRC.

The experimental results consistently demonstrated the superior performance of Deviation Networks compared to traditional methods. Deviation Networks exhibited higher accuracy, precision, recall, and F1 score, indicating their ability to effectively detect anomalies while minimizing false detections. The ROC curve and AUC values consistently showed that Deviation Networks outperformed traditional methods in capturing true anomalies while maintaining a low false positive rate. Similarly, the AUPRC values showcased the superiority of Deviation Networks in achieving high precision even with imbalanced datasets.

5.3 Quantitative Evaluation and Analysis:

In addition to the comparison with traditional methods, quantitative evaluations and analysis were performed to assess the performance of Deviation Networks in different scenarios. This analysis included studying the impact of varying anomaly ratios, exploring the robustness of Deviation Networks against noise and outliers, and investigating the detection capabilities in dynamic or evolving data streams.

The quantitative evaluations confirmed the robustness and adaptability of Deviation Networks across various scenarios. Deviation Networks showed consistent performance even with varying anomaly ratios, demonstrating their ability to handle imbalanced datasets effectively. They also exhibited resilience to noise and outliers, highlighting their capability to capture meaningful deviations and disregard irrelevant variations. Moreover, Deviation Networks showcased promising results in detecting anomalies in dynamic or evolving data streams, indicating their potential for real-time anomaly detection applications[4].

Overall, the experimental results on benchmark datasets and quantitative evaluations support the effectiveness and superiority of Deviation Networks in anomaly detection tasks. The next section will delve into practical considerations when implementing Deviation Networks, including data preprocessing, hyperparameter tuning, and deployment strategies[5].

6.Practical Considerations

Implementing Deviation Networks for real-world anomaly detection tasks involves several practical considerations to ensure optimal performance and successful deployment. This section discusses key considerations, including data preprocessing, hyperparameter tuning, and deployment strategies.

6.1 Data Preprocessing:

Data preprocessing plays a crucial role in anomaly detection tasks, as it helps in preparing the data for training Deviation Networks. Some important preprocessing steps include:

Data Cleaning: Remove any missing or inconsistent data points to ensure data quality.

Feature Selection/Extraction: Identify relevant features or perform feature engineering techniques to enhance the representation power of the data.

Normalization/Standardization: Scale the input data to a common range or transform it to have zero mean and unit variance, which can aid in network convergence and performance.

Handling Imbalanced Data: Address class imbalance issues, especially in datasets where anomalies are rare compared to normal instances. Techniques like oversampling, undersampling, or synthetic minority oversampling technique (SMOTE) can be applied to balance the dataset.

6.2 Hyperparameter Tuning:

Hyperparameter tuning is essential for optimizing the performance of Deviation Networks. Key hyperparameters to consider include:

Network Architecture: Experiment with different network architectures, such as the number of layers, layer sizes, and activation functions, to find the optimal configuration for the given dataset.

Learning Rate: Adjust the learning rate of the optimization algorithm to control the convergence speed and stability of the network.

Batch Size: Determine the batch size used during training, balancing computational efficiency and convergence quality.

Regularization Techniques: Apply regularization techniques like dropout or L2 regularization to prevent overfitting.

Reconstruction Error Threshold: Set an appropriate threshold for the reconstruction error to distinguish anomalies from normal instances. This threshold can be adjusted based on the desired trade-off between precision and recall.

Hyperparameter tuning can be performed using methods such as grid search, random search, or more sophisticated approaches like Bayesian optimization.

6.3 Deployment Strategies:

Once the Deviation Networks are trained and optimized, deploying them for real-world anomaly detection requires careful consideration:

Model Updates: Deviation Networks should be periodically retrained or updated to adapt to changing data patterns and new anomalies. This can involve retraining on a portion of the most recent data or employing online learning techniques to incorporate new data in real-time.

Scalability: Consider the scalability of Deviation Networks when dealing with large datasets or real-time streaming data. Techniques like distributed computing, parallel processing, or model compression can be employed to ensure efficient deployment.

Interpretability: Deviation Networks, being deep learning models, are often considered black-box models. Exploring techniques like model interpretability, feature importance analysis, or surrogate models can enhance the understanding and explainability of the anomaly detection results.

Integration with Existing Systems: Integrate Deviation Networks into existing systems or workflows to enable seamless anomaly detection. This may involve designing appropriate APIs, data pipelines, or incorporating the models into production environments.[6]

Considering these practical aspects can facilitate the successful implementation and deployment of Deviation Networks for real-world anomaly detection tasks.

7. Potential Applications and Use Cases:

Deviation Networks have a wide range of applications in various domains, including:

Cybersecurity: Detecting network intrusions, malware, or abnormal user behavior in computer networks.

Fraud Detection: Identifying fraudulent transactions or activities in financial systems.

Industrial Monitoring: Detecting anomalies in sensor data for predictive maintenance or quality control.Healthcare: Detecting abnormal patterns in patient health data for early disease diagnosis or monitoring.

IoT (Internet of Things): Anomaly detection in sensor data from IoT devices to ensure system integrity and security.

Image and Video Analysis: Identifying anomalies in images or videos for surveillance and security purposes.

Natural Language Processing: Detecting anomalies in text data.

7.Applications and Use Cases

eep Anomaly Detection with Deviation Networks has a wide range of applications across various domains. Here are some prominent use cases where Deviation Networks can be applied for anomaly detection.

7.1 Cybersecurity:

Deviation Networks can play a crucial role in detecting network intrusions, identifying abnormal user behavior, and identifying malware or malicious activities in computer networks. By analyzing network traffic patterns, system logs, and user behavior, Deviation Networks can effectively flag anomalies and potential security breaches, helping organizations enhance their cybersecurity defenses.

7.2 Fraud Detection:

In the financial sector, Deviation Networks can be utilized to detect fraudulent transactions or activities. By learning patterns from historical transaction data, Deviation Networks can identify deviations from normal transaction behaviour[7], thereby enabling timely detection and prevention of fraudulent activities such as credit card fraud, identity theft, or money laundering.

7.3 Industrial Monitoring:

Deviation Networks can be employed in industrial settings to monitor sensor data and detect anomalies in real-time. By analyzing sensor readings from machinery, equipment, or production processes, Deviation Networks can identify deviations that may indicate potential faults, breakdowns, or quality issues. This enables proactive maintenance, reduces downtime, and improves overall operational efficiency.

7.4 Healthcare:

In the healthcare domain, Deviation Networks can assist in early disease diagnosis, patient monitoring, and anomaly detection in medical imaging or physiological data. By analyzing electronic health records, vital signs, or medical images, Deviation Networks can flag abnormal patterns that may indicate the presence of diseases, adverse events, or anomalies requiring further investigation or intervention[9].

7.5 Internet of Things (IoT):

With the proliferation of IoT devices, Deviation Networks can be utilized to monitor and detect anomalies in sensor data from connected devices. Whether it's environmental sensors, smart grids, or industrial IoT deployments, Deviation Networks can analyze sensor data streams to identify abnormal patterns or anomalies, facilitating proactive maintenance, fault detection, and anomaly mitigation.

7.6 Image and Video Analysis:

Deviation Networks can be applied in image and video analysis for anomaly detection. They can identify unusual or suspicious activities in surveillance footage, detect objects or events that deviate from normal patterns, and assist in security and threat detection applications.

7.7 Natural Language Processing (NLP):

In the field of NLP, Deviation Networks can be utilized to detect anomalies in text data. This includes identifying unusual patterns or topics in social media posts, email communications, customer reviews, or news articles, which may indicate anomalies or emerging trends.

These are just a few examples of the diverse applications of Deviation Networks in anomaly detection. The adaptability and effectiveness of Deviation Networks make them suitable for various industries and domains where detecting anomalies and deviations is crucial for maintaining operational integrity, security, and efficiency.

8.Challenges and Future Directions

While Deep Anomaly Detection with Deviation Networks has shown promising results, there are still challenges and opportunities for further advancements in this field. Here are some key challenges and potential future directions:

8.1 Lack of Labeled Anomaly Data:

One of the primary challenges in anomaly detection is the scarcity of labeled anomaly data for training and evaluation. Collecting labeled anomalies can be time-consuming, expensive, or even impractical in some cases. Future research can focus on developing methods that require less labeled data or explore techniques for generating synthetic anomalies to augment the training process[10].

8.2 Handling High-Dimensional Data:

The Deviation Networks with Deep Anomaly Detection can face difficulties when dealing with high-dimensional data, such as images or text. The curse of dimensionality and the increased complexity of modelling high-dimensional data pose challenges in capturing meaningful deviations. Future research can explore techniques for dimensionality reduction, feature selection, or attention mechanisms specifically designed for high-dimensional data to enhance the performance of Deviation Networks.

8.3 Interpretable Anomaly Detection:

nterpretability is a critical aspect of anomaly detection, especially in domains where explanations for detected anomalies are required. Deep learning models, including Deviation Networks, are often considered black-box models, lacking interpretability. Future research can focus on developing techniques that provide interpretability and explanations for anomaly detection results, enabling users to understand the underlying factors contributing to anomalies.

8.4 Robustness to Adversarial Attacks:

Deep learning models, including Deviation Networks, are susceptible to adversarial attacks, where malicious entities intentionally manipulate data to evade anomaly detection. Ensuring the robustness and resilience of Deviation Networks against such attacks is an important research direction. Techniques like adversarial training, robust optimization, or anomaly detection methods specifically designed to handle adversarial attacks can be explored.

8.5 Streaming and Dynamic Data:

Real-time anomaly detection in streaming and dynamic data poses unique challenges. Deviation Networks need to adapt and update their models continuously to capture evolving anomalies. Future research can focus on developing online learning algorithms adaptive models, and efficient processing techniques to handle streaming data and enable real-time anomaly detection with Deviation Networks.

8.6 Transfer Learning and Domain Adaptation:

Applying Deviation Networks to new domains or datasets with limited labeled data can be challenging. Transfer learning and domain adaptation techniques can be explored to leverage pre-trained models or knowledge from related domains to enhance the performance of Deviation Networks. By transferring learned representations or adapting models to new domains, the need for extensive labeled data in every new domain can be alleviated.

8.7 Scalability and Efficiency:

As the scale and complexity of data continue to grow, scalability and efficiency become critical considerations. Future research can focus on developing scalable architectures, distributed computing techniques, and optimization algorithms to enable efficient training and deployment of Deviation Networks on large-scale datasets or in resource-constrained environments.

9. Conclusion

In conclusion, while Deep Anomaly Detection with Deviation Networks has demonstrated effectiveness in various applications, addressing challenges such as data scarcity, interpretability, robustness, and scalability will pave the way for future advancements. Overcoming these challenges and exploring new directions will further enhance the capabilities of Deviation Networks and enable their widespread adoption in anomaly detection tasks across diverse domains.

Deep Anomaly Detection with Deviation Networks has emerged as a powerful approach for detecting anomalies in various domains. By leveraging the representation learning capabilities of deep neural networks and the concept of deviation modeling, Deviation Networks have shown promising results in capturing anomalies and distinguishing them from normal instances.

In this paper, we explored the key components and techniques involved in Deep Anomaly Detection with Deviation Networks. We discussed the anomaly detection methods used for comparison, the architecture and functioning of Deviation Networks, the evaluation metrics employed to assess their performance, and the experimental results showcasing their superiority over traditional methods. We also delved into practical considerations, including data preprocessing, hyperparameter tuning, and deployment strategies.

Moreover, we highlighted the wide range of applications and use cases where Deviation Networks can be applied, such as cybersecurity, fraud detection, industrial monitoring, healthcare, IoT, image and video analysis, and natural language processing. These applications demonstrate the versatility and effectiveness of Deviation Networks in addressing real-world anomaly detection challenges.

However, we also acknowledged the challenges that need to be addressed and the potential future directions for further advancements in this field. These challenges include the scarcity of labeled anomaly data, handling high-dimensional data, interpretability, robustness to adversarial attacks, streaming and dynamic data, transfer learning, and scalability.

Despite these challenges, Deep Anomaly Detection with Deviation Networks holds great promise in revolutionizing anomaly detection tasks across diverse domains. With continued research and innovation, addressing the challenges and exploring new directions, Deviation Networks can further enhance their effectiveness, scalability, interpretability, and adaptability. This will enable their widespread adoption and contribute to improving the detection of anomalies, enhancing security, and optimizing various processes in real-world applications.

References :

[1] [1] Ruff, L., Vandermeulen, R., Goernitz, N., Deecke, L., Siddiqui, S. A., & Binder, A. (2018). Deep one-class classification. In International Conference on Machine Learning (pp. 4393-4402).

[2] 2] Zhou, C., Poria, S., Cambria, E., & Huang, G. B. (2017). Towards multimodal sentiment analysis: Harvesting opinions from the web. Data Mining and Knowledge Discovery, 31(6), 1673-1699.

[3] [3] Schlegl, T., Seeböck, P., Waldstein, S. M., Schmidt-Erfurth, U., & Langs, G. (2017). Unsupervised anomaly detection with generative adversarial networks to guide marker discovery. In International Conference on Information Processing in Medical Imaging (pp. 146-157).

[4] [4] Mahadevan, V., & Vasconcelos, N. (2010). Anomaly detection in crowded scenes. In IEEE Conference on Computer Vision and Pattern Recognition (pp. 1975-1981).

[5] [5] Chalapathy, R., & Chawla, S. (2019). Deep learning for anomaly detection: A review. ACM Computing Surveys, 51(3), 1-40.

[6] [6] Hodge, V. J., & Austin, J. (2004). A survey of outlier detection methodologies. Artificial Intelligence Review, 22(2), 85-126.

[7] [7] Liu, F. T., Ting, K. M., & Zhou, Z. H. (2012). Isolation forest. In Proceedings of the 2012 IEEE 12th International Conference on Data Mining (pp. 413-422).

[8] [8] Aggarwal, C. C. (2017). Outlier analysis. Springer.

[9] [9] Schölkopf, B., Platt, J. C., Shawe-Taylor, J., Smola, A. J., & Williamson, R. C. (2001). Estimating the support of a high-dimensional distribution. Neural Computation, 13(7), 1443-1471.

[10] [10] Ruff, L., & Vandermeulen, R. (2000). Deep one-class learning. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 10(4), e1352.