Anomaly Detection in Time Series and Spatial Data using Normalizing Flow

Chun Wei Hu (1226326896), Amarthya Kathula (1229583071)

December 7, 2024

Abstract

Anomaly detection is pivotal for autonomous systems and various real-world applications, ensuring operational safety and reliability. This research explores synthetic anomaly generation for spatial data and the application of machine learning techniques to detect anomalies in temporal datasets. Synthetic data generation leverages normalizing flows, such as Masked Autoregressive Flows (MAFs), to produce realistic anomaly patches for training segmentation models. This methodology supports the segmentation of anomalies as distinct classes, enhancing the robustness of anomaly detection in autonomous systems. Complementarily, machine learning methods like Isolation Forest, One-Class SVM, Autoencoders, and Normalizing Flows were employed to detect anomalies in temporal sensor data from a vehicle dataset. Among these, Normalizing Flow offered superior flexibility and interpretability, especially for complex, multimodal distributions, while Isolation Forest balanced efficiency and performance for simpler cases. Together, these approaches highlight the synergy of spatial and temporal anomaly detection techniques, paving the way for robust anomaly identification in diverse domains.

1 Introduction

Anomaly detection is crucial for many tasks, mainly for autonomous systems that operate in the human world, to avoid accidents. There is a good amount of research in this area, especially in image and time series anomaly detection [3] [6] [5]. But there is a lack of proper negative data with anomalies that can be used for research. The research we have done is on anomaly detection in spatial data for autonomous systems by generating synthetic images which aims at generating synthetic negative data for anomaly detection. It is mainly inspired from [6]. There are papers published on synthetic data generation, but it is difficult to find any open-source code that can be used by other researchers, so we aim to publish this on open-source platforms so that it can help others to focus more time on methodologies instead of trying to find negative data.

For spatial data, synthetic anomaly generation is explored using generative models such as Normalizing Flows. This approach supports the training of segmentation models for anomaly identification in image data, enabling autonomous systems to recognize out-of-distribution objects. For temporal data, machine learning techniques are employed to detect anomalies in time-series data derived from a vehicle's sensor readings, representing unusual driving behaviors or mechanical faults. Normalizing Flow offers several advantages, particularly its ability to provide explicit probabilistic scoring for anomalies, making it highly interpretable. The method is highly flexible and can model complex, multi-modal relationships that simpler models cannot. Additionally, its invertible architecture allows for inspecting how data is transformed into the base distribution. However, these benefits come with trade-offs. Normalizing Flow is computationally intensive, requiring significant resources for training, particularly on large datasets. It is also sensitive to hyperparameters, requiring careful tuning of architecture and coupling layers. Moreover, it assumes the availability of a sufficiently large dataset for effective training.

2 BackGround

2.1 Anomaly detection

Anomaly detection can be considered as a classification task which classifies out of distribution data from inlier data. This plays a crucial role in autonomous systems and is used in many real-life situations. There are handful of categories anomalies in autonomous systems can be classified into, such as based on camera data, lidar data, radar data, also multimodal data [1]. These classifications are made of either temporal or spatial or a mix of these two. We worked on spatial data, i.e., images taken from a driver perspective in a car [2] and a sample of temporal featured time series data.

2.2 Generative Models for Synthetic Data

Generative models are powerful tools for learning the underlying data distribution of inliers (normal data) and are widely used in anomaly detection. These models can be used to estimate the likelihood of new data points and generate synthetic anomalies by sampling from low-likelihood regions of the learned distribution. Popular generative models include Variational Autoencoders (VAEs) and Generative Adversarial Networks (GANs), which are effective but have limitations like approximate density estimation or training instability.

Normalizing flows, such as Masked Autoregressive Flows (MAFs), are particularly suited for this project because they provide exact likelihood estimation, making them ideal for density-based anomaly detection. Additionally, their invertible transformations enable efficient sampling and direct computation of probabilities, offering an advantage over VAEs and GANs.

For the generative model used for synthetic negative data generation, we used a normalizing flow-based generative model initialized with random weights. Then the MAF is trained to model the inlier distribution which are the extracted patches from the images in dataset and generate synthetic anomaly patches from its low-likelihood regions. For this implementation, we utilize the open source nflows library [4], which provides a robust framework for constructing and training normalizing flows. The MAF architecture is based on [4], which comprises components including masked transformers which are used to enforce the autoregressive property.

2.3 Segmentation Models

Segmentation models are essential for semantic segmentation tasks, where the goal is to assign a class label to every pixel in an image. These models are particularly relevant for anomaly detection, as they allow the identification of anomalous regions alongside normal classes. By training segmentation models to classify anomalies as a separate class in addition to standard semantic segmentation tasks, they become more robust in real-world scenarios, where out-of-distribution (OOD) objects or regions need to be identified accurately. We describe some state-of-the-art models used commonly in the related work section. The transformation models the conditional distribution of each variable given the previous ones, which is a hallmark of autoregressive models. This allows the MAF to capture complex dependencies in the data.

3 Related Work

The work on spatial anomaly detection was initially inspired from [5] which uses Jensen-Shannon divergence as a criterion of choice for robust joint learning in presence of noisy synthetic negatives which are generated by a jointly trained normalizing flow along with a discriminative model. Our initial plan was to use the same architecture as this but due to the complexity and required computational resources we decided to use a pre-trained segmentation model and fine tune it. They also provide evidence that coverage oriented synthetic negatives has with respect to their adversarial counterparts.

Also [7] provides, a comprehensive introduction to normalizing flows, including autoregressive flows like MAF explaining their ability to model complex probability distributions by learning bijective transformations which was helpful for this work and what inspired us to experiment with MAF for synthetic anomaly geeneration.

Last but no least, anomaly detection in time series data was inspired by [8] where an anomaly detection technique where unseen data samples are classified as normal or abnormal by scoring them against a learned model of normal data.

4 Technical Contribution

4.1 Spatial Data

We train a Masked Autoregressive Flow to generate synthetic images, and the proposed architecture involves fine tuning a segmentation model with the generated negative data to be able to detect anomalies. Architecture of the proposed model The proposed architecture involves training the model with mixed-content images which are generated by pasting a patch of random size which is synthetically generated using a generative model. This approach is mostly inspired from [6] where they use normalizing flows jointly trained with a segmentation model. From the images in data set we extract patches on which the MAF model is trained, and these patches are replaced with patches essentially images which are samples from the trained MAF. Instead of training the MAF separately we sample the patches from the beginning even when the MAF is not trained yet. One synthetic patch is samples and pasted into the inlier image which is then used to train the segmentation model. For training the MAF we use negative log likelihood of the samples as the loss function for estimating the true distribution of the patches. It involves training the segmentation model on these images, and the loss function used for this is a combination of segmentation loss (Standard cross-entropy for inlier regions) and Anomaly loss which encourages the model to classify synthetic anomalies as a new anomaly class.

4.2 Temporal Data

The analysis began with a key focus of this analysis was on Normalizing Flow, a deep probabilistic model that transforms complex data distributions into simpler ones, such as a Gaussian distribution, through a series of invertible transformations. Using RealNVP as the coupling layer architecture, Normalizing Flow applied scaling and shifting operations to map data into the base Gaussian distribution. Anomalies were detected based on low likelihood values derived from the negative log-likelihood. This method provides explicit probabilistic anomaly scoring, which is highly interpretable. However, its computational cost is higher compared to simpler methods, and the model is sensitive to hyperparameters, architecture design, and data preprocessing. Despite these challenges, Normalizing Flow offers flexibility and excels in datasets with intricate, multi-modal relationships. The detected anomalies is shown in Figure 1

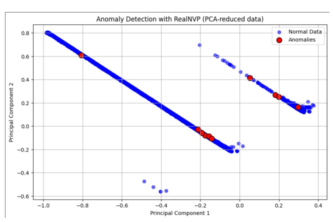

Figure 1: Normalizing Flow RealNVP Chart.

Next method will be Isolation Forest, a tree-based ensemble method that isolates anomalies by recursively partitioning the feature space. The method assumes that anomalies are data points that can be isolated quickly due to their sparsity or distinctiveness. Using 100 estimators and a contamination parameter of 0.05 (indicating 5Forest calculated anomaly scores based on the depth required to isolate each point. While this method is robust to high-dimensional data and does not assume any particular data distribution, it can struggle with irrelevant features and is sensitive to the contamination parameter. Although the chart looks like lots of anomalies being detected, but as shown in Figure 3, the correct amount of anomalies were being detected. The detected anomalies and scores are shown in Figure 2 and Figure 3.

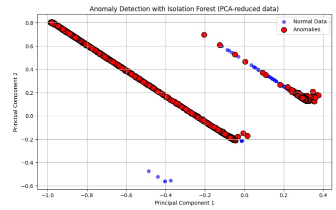

Figure 2: Isolation Forest Chart.

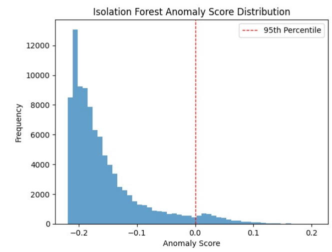

Figure 3: Isolation Forest Score.

Next, the One-Class SVM was applied, which uses a kernel-based approach to define a boundary around normal data points. Any point outside this boundary is classified as an anomaly. The model employed an RBF kernel, with a gamma value of 0.1 and a nu parameter of 0.05 to represent the expected anomaly proportion. This method effectively captures non-linear relationships in the data but is computationally expensive for large datasets and highly sensitive to hyperparameter tuning. The detected anomalies is shown in Figure 4

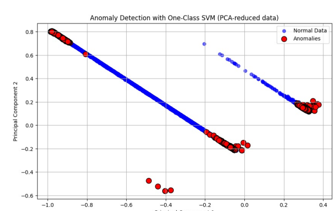

Figure 4: One-Class SVM Chart.

The Autoencoder, a neural network-based technique, was employed to reconstruct the input data. The model architecture consisted of two hidden layers with dropout regularization and a bottleneck of size 2. Anomalies were identified based on high reconstruction errors, as they deviated significantly from the learned patterns of normal data. The Autoencoder was trained over 100 epochs using a Mean Squared Error (MSE) loss function. While the Autoencoder excelled in capturing complex non-linear relationships in the data, it required careful architecture design and was prone to overfitting when anomalies dominated the dataset. The detected anomalies is shown in Figure 5 and error score is shown in Figure 6.

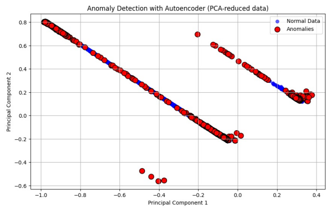

Figure 5: AutoEncoder Chart.

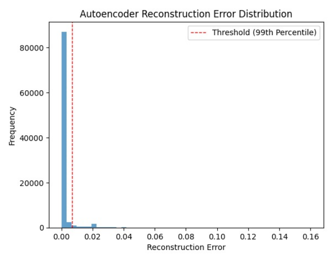

Figure 6: AutoEncoder Chart.

Lastly, the Elliptic Envelope method was applied. This method assumes that the data follows a multivariate Gaussian distribution and fits an ellipse around the central data points. Points outside this ellipse are classified as anomalies. With a contamination parameter of 0.05, the method used robust covariance estimation to handle noise and outliers. While it is simple and interpretable, its assumption of Gaussian data limits its applicability to real-world datasets with complex distributions. The detected anomalies is shown in Figure 7

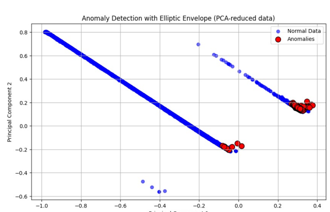

Figure 7: Elliptic Chart.

5 Evaluation and Experiment

5.1 Temporal Data

Each method was evaluated using standard metrics such as Precision, Recall, F1-Score, and ROC-AUC. For

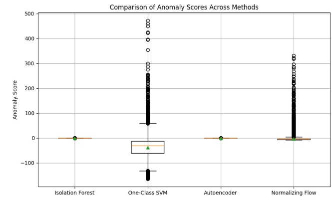

Figure 8: Comparison between different methods.

Isolation Forest, the model identified anomalies effectively with moderate ROC-AUC scores, provided the contamination parameter was appropriately set. One-Class SVMdemonstrated strong performance in capturing non-linear relationships but required substantial computational resources and careful tuning. The Autoencoder performed well for high-dimensional data, but overfitting remained a concern, as reconstruction errors were low for both normal and anomalous data points. Normalizing Flow delivered excellent results for complex data distributions, with probabilistic scoring providing additional interpretability, though it required significantly more computational effort. The Elliptic Envelope method, while computationally efficient, struggled with non-Gaussian distributions, making it less effective for this dataset. The evaluation scores are shown in Table 1.

	Group I		Group II	
	Precision	Recall	F1-Score	Roc-Auc
Isolation	0.0460	0.0474	0.0467	0.5006
One-Class	0.0441	0.0454	0.0447	0.5027
AutoEn	0.0451	0.0465	0.0458	0.5005
Normalize F	0.0490	0.0504	0.0497	0.5033
Elliptic	0.0553	0.0570	0.0561	0.4988

Table 1: Evaluation Scores

5.2 Spatial Data

The dataset used for this project is Cityascpaes [] which consists of 5000 images [3475 train and val sets, 1525 test images], each with a resolution of 1024 x 2048. We trained the MAF by using the train and val images.

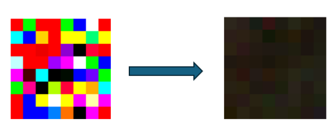

Figure 9: Synthetic patches generated by the MAF model before and after training.

The generates synthetic patches from the trained

MAF model can be seen the 9. The first image is from the standard distribution and the next image is samples from the MAF trained model trained on the images. For computational efficiency we resized the images to 512 x 512, but when the generated patches of random sizes are pasted onto the full resolution images, It is possible to generate anomalies which are more similar to real world anomalies.

Figure 10: Images before and after pasting the anomalies.

We were able to generate synthetic negative images with random sized patches on it which are shown in 11 by pasting the generated patches onto the images. And these images should be used to train the segmentation model.

Figure 11: Segmented Image using DeepLabV3 + Resnet101.

The fig 2 shows the segmentation model's results before attaching the synthetic patch. The model used for semantic segmentation of the images is DeepLabV3 + resnet101 [], trained on 768 x 768 randomly cropped images.

We were unable to fully implement the proposed architecture, which involved fine-tuning a pre-trained segmentation model on the generated images. However, we successfully achieved the synthetic generation of anomalous images.

6 Conclusion

This project focuses on anomaly detection in both spatial and temporal featured data. From the work done with temporal features we can conclude that for datasets with simple distributions or minimal features, methods such as Isolation Forest or Elliptic Envelope are sufficient due to their simplicity and computational efficiency. However, for complex datasets with intricate relationships, Normalizing Flow or One-Class SVM methods are more appropriate. While Normalizing Flow outperforms other methods in flexibility and interpretability, its computational cost and sensitivity to configuration should be carefully considered. For real-time or large-scale applications, Isolation Forest remains a practical choice, balancing efficiency and performance.

From the work done with spatial features we conclude that the MAF successfully generated patches that were visually plausible and varied, showcasing its potential to model the intricacies of the Cityscapes dataset. Sampling methods for learning the true data distribution play a critical role in the effectiveness of generative models, rejection sampling while effective for filtering samples closer to inliers, is computationally expensive and may not be the most efficient approach in practice. However, fine-tuning the sampling strategy to consistently generate "boundary anomalies" (slightly outside the inlier distribution) required additional experimentation.

7 Future Work

In the future we are interested in implementing the anomaly detection for images using the proposed synthetic anomaly generation and also use more refined and feasible sampling techniques.

We also observed that simple normalizing flows has it's own advantages and we'd like to weigh the differences between these two approaches which we intended to do but couldn't due restrictions in time. We aim to open source a fully functional implementation of the anomaly detection in spatial features area as it can help others avoid the problem of finding synthetic data generating approaches or building one from scratch.

References

[1] Daniel Bogdoll, Maximilian Nitsche, and J. Marius Zollner. Anomaly detection in autonomous driving: A survey. In *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, page 4487–4498. IEEE, June 2022.

[2] Marius Cordts, Mohamed Omran, Sebastian Ramos, Timo Rehfeld, Markus Enzweiler, Rodrigo Benenson, Uwe Franke, Stefan Roth, and Bernt Schiele. The cityscapes dataset for semantic urban scene understanding. In *Proc. of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016.

[3] Anja Delić, Matej Grcić, and Siniša Šegvić. Real time dense anomaly detection by learning on synthetic negative data, 2023.

[4] Conor Durkan, Artur Bekasov, Iain Murray, and George Papamakarios. nflows: normalizing flows in PyTorch, November 2020.

[5] Matej Grcić, Petra Bevandić, Zoran Kalafatić, and Siniša Šegvić. Dense out-of-distribution detection by robust learning on synthetic negative data, 2023.

[6] Matej Grcić, Petra Bevandić, and Siniša Šegvić. Densehybrid: Hybrid anomaly detection for dense open-set recognition, 2022.

[7] George Papamakarios, Eric Nalisnick, Danilo Jimenez Rezende, Shakir Mohamed, and Balaji Lakshminarayanan. Normalizing flows for probabilistic modeling and inference, 2021.

[8] Maximilian Schmidt and Marko Simic. Normalizing flows for novelty detection in industrial time series data, 2019.