How Reliable are the Deep Learning-based Anomaly Detectors? A Comprehensive Reliability Analysis of Autoencoder-based Anomaly Detectors

Manzoor Hussain, Jae-Won Suh, Bo-Seok Seo, Jang-Eui Hong*
College of Electrical & Computer Engineering
Chungbuk National University
Cheongju, Republic of Korea
{manzoorhussain, boseok, sjwon, jehong}@chungbuk.ac.kr

Abstract**—** Autoencoders have been successfully used for detecting anomalies in various intelligent systems. Generally, the autoencoder-based anomaly detection system detects a data point as anomalous if its reconstruction loss is higher than the set threshold. However, the autoencoder-based anomaly detection system lacks generalizability and reliability. Theoretically, autoencoders should produce reconstruction loss higher than the set threshold for all out-of-distribution data points. However, it has been noted that the autoencoders produce lower reconstruction loss for some data points and higher for some specific data points that are out of distribution. For example, autoencoders produce lower reconstruction loss than the set threshold for the input image with motion blur perturbation. In contrast, it produces a higher reconstruction loss for the input image with impulse noise. Motivated by the above-mentioned problem statement, this paper aims to investigate the autoencoders' reliability in detecting various anomalies. Our experimental results found that autoencoders are highly data-specific, and the detection results are highly unreliable for various anomalies. We proposed a framework with designed reliability evaluation metrics to help assess the reliability of autoencoder-based anomaly detectors. We reported the analysis of the reliability of autoencoders in terms of their reconstruction loss against perturbed versus normal images.

Keywords**— *Anomaly Detection, Reliability, Autoencoders, Deep Learning, Safety***

I. Introduction

An Anomaly or outlier is a data point that significantly differs from other data points in the n-dimensional dataset and does not conform to normal behavior. Any deviation in the established baseline pattern in the dataset is considered an anomalous data point or an outlier. Anomaly detection is a process of identifying novelty or abnormal data points in the dataset[1]. It can be regarded as identifying the abnormal distribution of data points that do not exhibit the expected distribution that should be consistent with the overall dataset. Generally, the autoencoder-based anomaly detection system differentiates whether a data point is normal or abnormal based on the reconstruction error. A reconstruction error (i.e., often calculated in mean squared error (MSE)) is the difference between the input and reconstructed output. Typically, the autoencoders, such as variational autoencoders, convolutional autoencoders (CAE), and deep autoencoders (DAE), are first trained with the normal dataset to minimize the reconstruction errors. Once these autoencoders are trained, they learn to reconstruct the normal data points from their low dimensional latent features; however, they do not reconstruct well in the case of abnormal data points. Additionally, the anomalous data points produce larger reconstruction errors, while on the other hand, normal data points produce lower reconstruction errors. Thus, by comparing the reconstruction errors against the suitable threshold, the anomaly detection system distinguishes whether the data points are anomalous or normal. However, sometimes the autoencoders suffer from a lack of stability in their performance in detecting anomalous data points.

For example, a variational autoencoder was trained with the normal images frame without abnormal images to detect anomalous images that contain *speckle noise, impulse noise, pixelating,* and *motion blurring* in vision-based autonomous driving systems. Theoretically, the trained autoencoders should produce higher reconstruction for the anomalous images containing *speckle noise, impulse noise, pixelate,* and *motion blurring* than the normal images. However, the autoencoder output differed for the different image perturbation types during the experiment. The reconstruction errors for *motion blurring* and *pixelate* perturbation was less than the normal image. As a result, image frames with *motion blurring* and *pixelate* perturbation were not detected as anomalous data points. On the other hand, the DNN model in the vision-based autonomous driving system produces an incorrect output which causes the Autonomous Driving System (ADS) to start violating the lane-keeping safety requirement. Ideally, the autoencoder reconstruction errors should be higher for all image perturbation types than for the normal image. However, in practice, we found the autoencoders are data-specific [2], meaning that their performance in detecting anomalies varies as per the nature of the data. Therefore, it is hard to rely on an autoencoder-based anomaly detection system to detect anomalies for safety-critical systems like vision-based ADSs.

Motivated by the above-mentioned problem statement, we performed a comprehensive reliability analysis of autoencoder-based anomaly detection systems. To justify our claim regarding the reliability of autoencoders, a case study on anomaly detection based on auto-encoders in vision-based ADSs was conducted and analyzed its reliability. Our experiments' results will help guide the selection of suitable

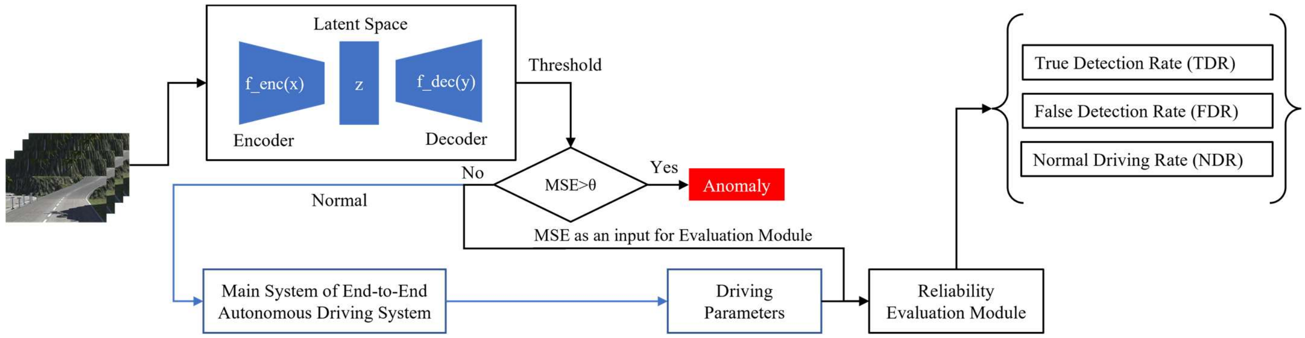

Figure 1. The proposed approach for reliability analysis of autoencoders-based anomaly detection in vision-based autonomous driving systems.

autoencoders to detect different types of anomalies in the image-based dataset. We made the following contributions to this article.

1. We propose a reliability analysis framework for the autoencoder-based anomaly detection systems that directly take the produced reconstruction loss of autoencoders as input and produce the score for reliability metrics.
2. We conducted extensive experiments, observed the reconstruction loss of autoencoders for different anomalous inputs (e.g., *speckle noise, impulse noise, pixelate, motion blur, etc.*), and evaluated the reliability score.
3. We proposed reliability analysis metrics that can be used to evaluate whether or not the autoencoders are reliable for anomaly detection based on the reconstruction loss produced for various kinds of anomalous inputs.

The rest of the article is organized as follows. Section II introduces some studies related to our work and briefly explains the proposed approach in Section III. Section IV presents the experimental setup and results. Finally, we conclude our work in Section V.

II. Related Work

Anomaly detection has been an extensively studied field since 1960. It has been used in a wide range of practical applications, including fraud detection[3], [4], detecting abnormality in autonomous driving behavior [5]–[8], network security[9], medical diagnosis [10]–[12], and multi-sensor data [13]. Data-driven anomaly detection is becoming more popular due to the availability of enormous amounts of data from sensors. Therefore, the application of deep learning and neural network-based anomaly detection significantly influenced the field of anomaly detection. Deep learning methods have shown excellent performance compared to traditional anomaly detection methods [14].

Various research works have explored the flaws in autoencoder-based anomaly detection systems, such as the authors [15] described that the autoencoders are highly dependent on training data. It may struggle to detect anomalies that deviate considerably from training data. In another study, the authors [16] explored autoencoders' limitations regarding their performance on high-dimensional data. According to the study, the autoencoders may struggle to detect anomalies in high dimensional data such as image or time series data with a large number of features. Zhou et al. [17] studied the robustness of autoencoders for anomaly detection. In contrast to existing studies, we studied the reliability of autoencoders in this article and proposed evaluation metrics to assess the reliability.

III. Proposed Approach

In recent studies, autoencoder-based anomaly detectors in various domains have shown excellent results among unsupervised learning-based methods. Many researchers used various kinds of autoencoders to detect anomalies and outliers, such as Salvic, Guilia, et al. [18] used variational autoencoders to detect multilevel anomalies. At the same time, Stocco et al. [5] used DAE, CAE, variational autoencoders (VAE), and long-short-term memory autoencoders to detect misbehavior in autonomous vehicles. Esma Mujkic et al. [19] used autoencoders to detect anomalies in agricultural vehicles.

Besides the autoencoder's great success in detecting an anomaly, studies have shown that it has some stability flaws which affect the overall performance and reliability of anomaly detection results. Therefore, we propose a reliability evaluation method that can be used to check whether or not the state-of-the-art autoencoder-based anomaly detection systems are reliable. In order to validate our proposed approach, we considered autoencoder-based anomaly detection used in vision-based ADSs and evaluated its reliability using the proposed approach. We first trained the anomaly detector and the DNN model used in ADSs with normal images in our proposed approach. Then we validate the performance of autoencoders in detecting the anomalous image streams. To create the anomalous image stream, we applied basic image corruption techniques such as *speckle noise, impulse noise, pixelate, motion blurring*, etc. The evaluation metrics take the DNN model's output and the MSE of the autoencoder-based anomaly detector as input to produce the score for metrics, in order to perform reliability analysis. Figure 1 shows the overall working flow of the proposed approach. In the following subsection, we discuss the core module in detail.

A. Autoencoders-based Anomaly Detectors

Autoencoders are a special neural network designed to learn a compressed representation of the input data. It has two main components. The encoding part is responsible for mapping the input data into lower-dimensional encoding. On the other hand, the decoding part maps the encoded data back to the original input data. The autoencoders are trained by minimizing the reconstruction loss, which is the actual metric used to measure the difference between the original and reconstructed input. Let us consider that we have an image stream *I* and an encoding function $z = f\_enc(I)$ that is used to encode the input image stream *I*. The decoder part of the autoencoder can be expressed as a function of $I_{rec} = f\_dec(z)$, where the I_{rec} is the reconstructed version of the original input image *I*. The difference between the original and reconstructed images is called the reconstruction loss, and it is used as an anomaly score to detect the anomalous input image. The general pixel-wise reconstruction losss in MSE can be calculated using equation 1.

$$MSE = \frac{1}{N \times H \times W \times C} \sum_{i=1,j=1,k=1,c=1}^{N,H,W,C} (I[i,j,k,c] - I'(i,j,k,c))^2 \qquad (1)$$

Where *N* is the number of samples, H represents the height, *W* represents the width, and *C* represents the channel. Note that in this experiment, we evaluated the reliability of the CAE, VAE, and DAE. These autoencoders are the most commonly used anomaly detectors.

B. Reliability Evaluation Metrics

An autoencoder-based anomaly detector is considered reliable if it successfully detects all anomalous streams as anomalous. While on the other hand, an autoencoder-based anomaly detector can be considered an unreliable anomaly detector if it does not detect an anomalous image stream as an anomaly. This unreliable detection results of an anomaly can cause certain deviations in the final output of the DNN model used in vision-based ADSs. The definition of reliability in the context of this article is that the *autoencoder-based anomaly detection system should perform its intended function without any false detection of anomalous data points*. The evaluation metrics take the final output of the ADSs model and the reconstruction loss of autoencoders, which is calculated in MSE to determine their values. We proposed three metrics: True Detection Rate (TDR), False Detection Rate (FDR), and Normal Driving Rate (NDR). This experiment was conducted using the Udacity self-driving car simulator, and the NVIDIA DAVE-2 model was trained using normal images for lane-keeping purposes. Note that the NVIDIA DAVE-2 model directly maps the input image coming from the front camera of ADSs into driving parameters. We used three types of autoencoders: CAE, VAE, and DAE for anomaly detection purposes. We then performed the reliability analysis of each autoencoder and reported the results in terms of reconstruction loss using proposed metrics. The proposed evaluation metrics for reliability analysis are defined as follows:

True Detection Rate (TDR): We defined TDR as the rate of detecting anomalous image streams correctly by an anomaly detector. The value of TDR is defined based on the reconstruction loss of autoencoders. When the reconstruction loss is greater than the threshold for the anomalous image streams, the anomalous image streams are identified as anomalous. Subsequently, if the autoencoder-based anomaly detector detects an anomalous stream successfully, the value of TDR is increased by one. We can define the TDR as follows:

$$TDR = I'/k_{dts} \qquad (2)$$

Where k_{dts} is the dataset that contains normal and anomalous images and I' is the total number of anomalous image streams that were successfully detected by the anomaly detector.

False Detection Rate (FDR): The FDR is the anomaly detector's rate of undetected anomalous image streams. The value of FDR is determined by the reconstruction loss of autoencoders as well as the deviation of the final output of the DNN model. Consider that we have an anomalous image stream I', MSE denotes the reconstruction loss of the anomaly detector, and the threshold is denoted by θ. Let us assume that we have a DNN model $y = f(O, I)$, where the O is the actual output of the model $f(O, I)$ for the normal image *I*. The anomaly detector aims to detect the anomalous image so that the model should not produce incorrect output O'. Thus, we can determine the value of FDR based on two conditions. Firstly, the output of the model $f(O, I')$ for all anomalous streams I' should deviate from its original output *O*. Secondly, in contrast to TDR, the reconstruction loss should be less than the θ. This means that the reconstruction loss of the image stream was less than the threshold. Therefore, it could not be detected as anomalous streams. Consequently, the DNN model used in vision-based ADSs produces incorrect outputs, violating the safety requirement (i.e., lane keeping).

Normal Driving Rate (NDR): We define the NDR as the percentage of safe trajectories in total trajectories produced by the DNN model $f(O, I)$ for all image streams in the dataset k_{dts}. These are the trajectories where no violation of safety requirements has occurred. In other words, it is the percentage of normal driving behavior where the ADSs system follows the center of the road whether or not the input image stream is normal or anomalous. This value is determined by the reconstruction loss as well as the output of the DNN model used in vision-based ADSs.

IV. Experimental Setup and Results

A. Experimental Setup

We conducted a brief experiment in order to evaluate the reliability of the commonly used autoencoders for anomaly

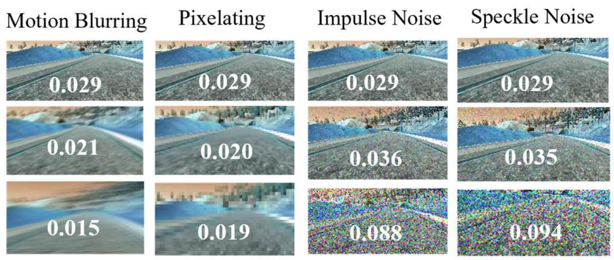

Figure 2. Reconstruction loss for normal image frames and corrupted image. The first row represents the normal image frames and their reconstruction loss, while the second row represents the low-level image corruption applied and their reconstruction. Finally, the last row represents the higher level of image corruption applied to the image frames and their reconstruction loss. The sample was collected when VAE was used as an anomaly detector.

detection. This study will help select reliable autoencoders that can be used as anomaly detectors in safety-critical applications such as ADSs. In our experiment, we use CAE, VAE, and DAE to detect anomalies in the driving behavior of the NVIDIA DAVE-2 ADSs model as a case study using Udacity self-driving car simulator. We first trained the autoencoder-based anomaly detector and NVIDA's model for lane-keeping purposes using normal images without anomalies and recorded the driving behavior. To assess the reliability of the anomaly detector, we applied common image corruption techniques to the input image, such as *speckle noise, impulse noise, pixelate,* and *motion blurring*. We set two levels for the intensity of image perturbation: The high and low levels. However, we found that even for the minor level of image corruption, there was a significant impact on anomaly detector reliability. We then evaluated the reliability of each autoencoder in detecting anomalies in the input image and reported the findings in the following subsection. A reliable anomaly detector should ideally flag the corrupt input image as an anomaly. On the other hand, a non-reliable anomaly detector may not flag the corrupted input image as an anomaly.

B. Results

In this experiment, the evaluation of the reliability of the anomaly detector is based on the reconstruction loss of the autoencoders and the trajectories produced by the model used in vision-based ADSs. The impact of the different types of perturbations and their reconstruction loss is shown in figure 2. The threshold for the anomaly detector was set to 0.020. This means the autoencoder should flag an image frame as anomalous if the reconstruction loss is higher than 0.020. However, in our experiment, surprisingly, we found that the autoencoders-based anomaly detector performed well in detecting some specific types of anomalies, such as impulse noise and speckle noise, while on the other hand, it performed worst in detecting other types of anomalies, such as motion blur and pixelating. Interestingly, instead of increasing the reconstruction loss for increased perturbation, the reconstruction loss decreased when the intensity was increased for both motion blurring and pixelating. As a result, the anomaly detector could not flag those image streams as an anomaly. Consequently, the FDR increased when pixelating and motion blurring intensity increased. In contrast to pixelating and motion blurring, the autoencoder's reconstruction loss was quite high for low-level and high-level impulse and speckling noise perturbation. As a result, the anomaly detector successfully flagged those image frames as anomalous. Hence the TDR for impulse and speckle noise increased significantly.

We present the effect of pixelating, motion blurring, impulse noise, and speckle noise on the reconstruction loss in figure 3. The numerical results of the reliability evaluation score are presented in Tables 1 and 2.

Table 1. Numerical results of reliability metrics under high-intensity perturbation.

Autoencoders	Anomaly Types	TDR	FDR	NDR
VAE	Impulse noise	0.977	0.022	0.00
	Motion blur	0.000	1.000	0.00
	Pixelate	0.066	0.933	0.00
	Speckle Noise	0.977	0.022	0.066
CAE	Impulse noise	0.979	0.020	0.020
	Motion blur	0.040	0.959	0.00
	Pixelate	0.081	0.918	0.00
	Speckle Noise	0.976	0.020	0.00
DAE	Impulse noise	0.976	0.023	0.046
	Motion blur	0.000	1.000	0.00
	Pixelate	0.069	0.930	0.00
	Speckle Noise	0.976	0.023	0.00

Table 2. Numerical results of reliability metrics under low-intensity perturbation.

Autoencoders	Anomaly Types	TDR	FDR	NDR
VAE	Impulse noise	0.991	0.008	0.008
	Motion blur	0.973	0.026	0.062
	Pixelate	0.991	0.008	0.116
	Speckle Noise	0.991	0.008	0.008
CAE	Impulse noise	0.976	0.020	0.040
	Motion blur	0.976	0.020	0.102
	Pixelate	0.976	0.020	0.000
	Speckle Noise	0.976	0.020	0.081
DAE	Impulse noise	0.991	0.008	0.025
	Motion blur	0.991	0.008	0.042
	Pixelate	0.991	0.008	0.128
	Speckle Noise	0.991	0.008	0.051

Figure 3 shows that the autoencoder could successfully detected almost every speckle and impulse noise-affected anomalous image stream. However, in the case of motion blurring and pixelating type perturbation, the autoencoder shows highly unreliable and weird detection results. For example, when the intensity of the perturbation was set to a lower level, the autoencoders could detect almost every anomalous image stream. On the other hand, unexpectedly, when we set the perturbation level to the highest level, the reconstruction loss for all pixelating and motion-blurring anomalous streams was surprisingly very low. Additionally, all the autoencoders could not detect highly-level motion blurring and pixelating perturbation. In both cases, the output (i.e., Steering angle) of the DNN model strongly deviated from its normal output, violating the lane-keeping safety requirement. Consequently, the DFR increased drastically when motion blurring and pixelating intensity were set to the highest level. On the other hand, an autoencoder-based anomaly detector could detect almost all speckle and impulse-induced anomalous streams, and hence the TDR also increased.

Table 1 presented the evaluation score for each autoencoder when the perturbation level was set to the highest. While on the other hand, Table 2 presents the evaluation score for each autoencoder-based anomaly detector when the perturbation level was set to its lowest level. From Table 1, we can observe that when impulse and speckle noise was added to the input image frames, the TDR of all autoencoders was above 97% when the intensity was set to highest. Similarly, from Table 2, we can see that when the perturbation level of speckle and impulse noise was set to the lowest, the TDR of VAE and DAE was recorded at 99%. In comparison, the TDR of CAE was recorded at 97%. Conversely, we observe low FDR in both the highest and lowest level speckle and impulse noise perturbation. For example, we can see that the FDR of VAE, CAE, and DAE was less than 2.3 % when we applied high-impulse noise to the image stream. The same trend was observed in the case of speckle noise, as we can observe in both Tables 1 and 2.

In contrast to speckle noise and impulse noise, all the autoencoder-based anomaly detectors produced unreliable detection results for motion blurring and pixelated perturbation. From both tables, we can see that the detection result is highly unreliable. The FDR is quite high (i.e., in some cases, it is 100%) for both motion blurring and pixelating. An interesting phenomenon was observed when the intensity was higher. The reconstruction loss for both motion-blurring and pixelating perturbation was less than the normal images when we applied higher-level pixelating and motion-blurring perturbation. Which, consequently, increases the FDR. Therefore, in the case of motion blurring and pixelating, the FDR of all autoencoders is higher than the FDR of impulse and speckle noise. Table 1 shows that the FDR of VAE was 100% and 93%, respectively,

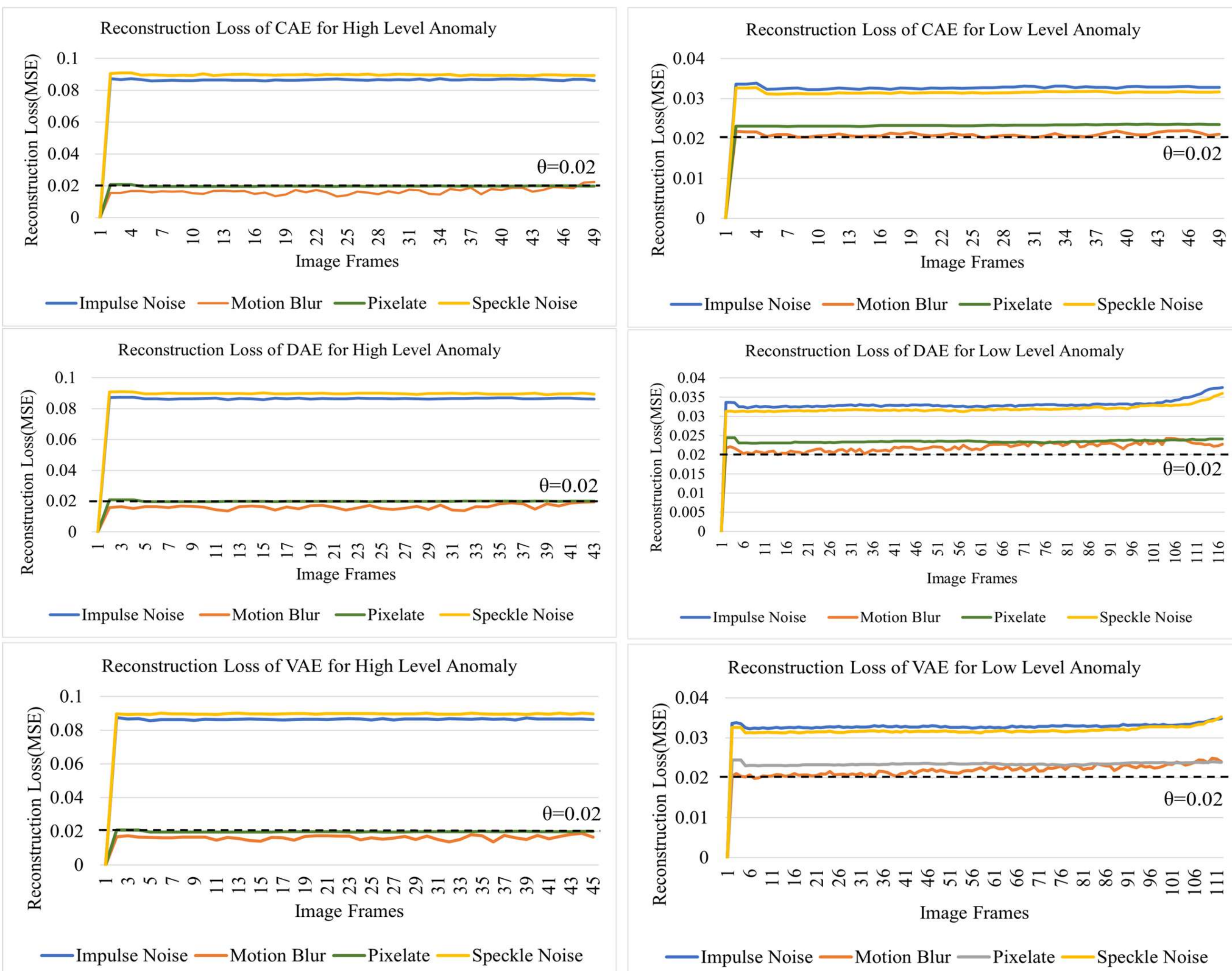

Figure 3. Reconstruction loss of autoencoders used for anomaly detection. The left column represents the reconstruction loss for high-intensity perturbation, while the right column represents the loss for low-intensity perturbation of VAE, DAE, and CAE.

and when the intensity of motion blurring and pixelating was high. Conversely, we observed lower FDR when motion blurring and pixelating intensity were lower listed in Table 2. Regarding the NDR, under all perturbations, the NDR was quite low. This means that the NVIDIA model could not understand the scene due to the anomalous image stream, so it does not follow the lane-keeping safety requirement.

V. CONCLUSION

We systematically investigated the reliability of deep learning-based anomaly detectors focusing on the autoencoders and presented a case study on the autoencoder-based anomaly detectors used in vision-based autonomous driving systems. Specifically, we proposed a reliability evaluation framework with metrics in order to perform a reliability assessment. The extensive experiment proved that autoencoder-based anomaly detectors are unreliable for specific types of anomalies, such as motion blurring and pixelating, while showing highly reliable detection results against specific types of an anomaly, such as speckle noise and impulse noise. This finding can help to select appropriate anomaly detectors for safety-critical systems.

In the future, we plan to expand our experiment to investigate the reason behind this phenomenon and propose a counterfactual explanation of reliable and unreliable detection results of deep learning-based anomaly detection systems.

ACKNOWLEDGMENT

This research was supported by the National Research Foundation of Korea (NRF) grant funded by the Ministry of Science and ICT. (NRF-2020R1A2C1007571).

REFERENCES

[1] S. Ryu, J. Yim, J. Seo, Y. Yu, and H. Seo, "Quantile Autoencoder with Abnormality Accumulation for Anomaly Detection of Multivariate Sensor Data," *IEEE Access*, vol. 10, pp. 70428–70439, 2022, doi: 10.1109/ACCESS.2022.3187426.

[2] A. Maheshwari, P. Mitra, and B. Sharma, "Autoencoder: Issues, Challenges and Future Prospect," *Lect. Notes Mech. Eng.*, pp. 257–266, 2022, doi: 10.1007/978-981-16-9236-9_24/COVER.

[3] P. Jiang, M. Eng, J. Zhang, (M Eng, and J. Zou, "Credit Card Fraud Detection Using Autoencoder Neural Network," 2509.

[4] M. Rezapour, "Anomaly detection using unsupervised methods: Credit card fraud case study," *Int. J. Adv. Comput. Sci. Appl.*, 2019, doi: 10.14569/IJACSA.2019.0101101.

[5] A. Stocco, M. Weiss, M. Calzana, and P. Tonella, "Misbehaviour prediction for autonomous driving systems," in *Proceedings of the ACM/IEEE 42nd International Conference on Software Engineering*, Jun. 2020, pp. 359–371, doi: 10.1145/3377811.3380353.

[6] N. Chakraborty *et al.*, "Structural Attention-Based Recurrent Variational Autoencoder for Highway Vehicle Anomaly Detection; Structural Attention-Based Recurrent Variational Autoencoder for Highway Vehicle Anomaly Detection," Accessed: Feb. 01, 2023. [Online]. Available: https://sites.google.com/illinois.edu/saber-vae.

[7] A. Stocco and P. Tonella, "Towards Anomaly Detectors that Learn Continuously; Towards Anomaly Detectors that Learn Continuously," *2020 IEEE Int. Symp. Softw. Reliab. Eng. Work.*, 2020, doi: 10.1109/ISSREW51248.2020.00073.

[8] M. Hussain, N. Ali, and J. E. Hong, "DeepGuard: a framework for safeguarding autonomous driving systems from inconsistent behaviour," *Autom. Softw. Eng.*, 2022, doi: 10.1007/s10515-021-00310-0.

[9] T. Zixu, K. S. K. Liyanage, and M. Gurusamy, "Generative Adversarial Network and Auto Encoder based Anomaly Detection in Distributed IoT Networks," 2020, doi: 10.1109/GLOBECOM42002.2020.9348244.

[10] T. Fernando, H. Gammulle, S. Denman, S. Sridharan, and C. Fookes, "Deep Learning for Medical Anomaly Detection A Survey," *ACM Comput. Surv.*, 2022, doi: 10.1145/3464423.

[11] P. Liu, X. Sun, Y. Han, Z. He, W. Zhang, and C. Wu, "Arrhythmia classification of LSTM autoencoder based on time series anomaly detection," *Biomed. Signal Process. Control*, 2022, doi: 10.1016/j.bspc.2021.103228.

[12] W. Gerych, E. Agu, and E. Rundensteiner, "Classifying Depression in Imbalanced Datasets Using an Autoencoder- Based Anomaly Detection Approach," 2019, doi: 10.1109/ICOSC.2019.8665535.

[13] L. Erhan *et al.*, "Smart anomaly detection in sensor systems: A multi-perspective review," *Inf. Fusion*, 2021, doi: 10.1016/j.inffus.2020.10.001.

[14] M. Mohammadi, A. Al-Fuqaha, S. Sorour, and M. Guizani, "Deep learning for IoT big data and streaming analytics: A survey," *IEEE Communications Surveys and Tutorials*. 2018, doi: 10.1109/COMST.2018.2844341.

[15] R. Chalapathy and S. Chawla, "Deep Learning for Anomaly Detection: A Survey," Jan. 2019, doi: 10.48550/arxiv.1901.03407.

[16] J. T. Springenberg, A. Dosovitskiy, T. Brox, and M. Riedmiller, "Striving for simplicity: The all convolutional net," 2015.

[17] C. Zhou and R. C. Paffenroth, "Anomaly detection with robust deep autoencoders," 2017, doi: 10.1145/3097983.3098052.

[18] G. Slavic, M. Baydoun, D. Campo, L. Marcenaro, and C. Regazzoni, "Multilevel Anomaly Detection Through Variational Autoencoders and Bayesian Models for Self-Aware Embodied Agents," *IEEE Trans. Multimed.*, 2022, doi: 10.1109/TMM.2021.3065232.

[19] E. Mujkic, M. P. Philipsen, T. B. Moeslund, M. P. Christiansen, and O. Ravn, "Anomaly Detection for Agricultural Vehicles Using Autoencoders," *Sensors 2022, Vol. 22, Page 3608*, vol. 22, no. 10, p. 3608, May 2022, doi: 10.3390/S22103608.