Leveraging Convolutional Neural Network (CNN)-based Auto Encoders for Enhanced Anomaly Detection in High-Dimensional Datasets

M. Aetsam Javed
Department of Computer Science, Faculty of Computer Science & IT, Superior University Lahore, Pakistan
aetsam.javed8287@gmail.com

Madiha Anjum
School of Information Technology, Victoria University, Australia
madiha.anjum@vu.edu.au

Hassan A. Ahmed
Information Systems, Cleveland State University, Ohio, USA
h.a.ahmed@csuohio.edu

Arshad Ali
Faculty of Computer and Information Systems, Islamic University of Madinah, Al Madinah Al Munawarah, Saudi Arabia
a.ali@iu.edu.sa

H. M. Shahzad
Faculty of Computer Science and Information Technology, Superior University Lahore, Pakistan
shahzad.dar@gmail.com

Hamayun Khan
Department of Computer Science, Faculty of Computer Science & IT, Superior University Lahore, Pakistan
hamayun.khan@superior.edu.pk (corresponding author)

Abdulaziz M. Alshahrani
Faculty of Computer and Information Systems, Islamic University of Madinah, Al Madinah Al Munawarah, Saudi Arabia
alshahrani@iu.edu.sa

Received: 8 August 2024 | Revised: 29 August 2024 and 20 September 2024 | Accepted: 22 September 2024

Licensed under a CC-BY 4.0 license | Copyright (c) by the authors | DOI: https://doi.org/10.48084/etasr.8619

ABSTRACT

This study presents an Auto-Encoder Convolutional Neural Network (AECNNs) approach for anomaly detection in high-dimensional datasets. Unsupervised learning-based algorithms have a strong theoretical foundation and are widely used for anomaly detection in high-dimensional datasets, but some limitations significantly reduce their performance. This study proposes an algorithm to address these limitations. The proposed AECNN combines various convolutional layers, feature extraction, dimensionality reduction, and data preprocessing and was evaluated using accuracy, precision, recall, and F1-score. The performance of

the proposed model was evaluated using a large real benchmark dataset. The proposed CNN-based autoencoder distinguished anomalies with an AUC score of 0.83 and remarkable accuracy, precision, recall, and F1 score.

Keywords-autoencoders; anomaly detection; high-dimensional data; machine learning; data analysis; model evaluation; Convolutional Neural Networks (CNNs); NSL-KDD; UNSW-NB15; MSE

I. INTRODUCTION

Anomaly detection is an important element of data examination that differentiates data that deviate from predicted behavior. These unusual irregularities, known as anomalies in a dataset, can indicate important events that require immediate attention, such as extortion in financial transactions, restorative conditions in healthcare diagnostics, and security breaches in cybersecurity frameworks. The primary objective of anomaly detection is to recognize unusual and possibly affected behavior from normal data [1]. Conventional procedures such as rule-based systems and measurable methods have been utilized for areas with high data insecurity to secure data over the network by avoiding anomalies to handle the complexity and high volume of financial data transactions. In [2], a PCA-based approach was proposed to reduce dimensionality in a network dataset. Unusual patterns in crucial healthcare data, such as vital signs and symptomatic images, should be quickly identified. In any case, the complex nature of clinical records and the requirement for high accuracy show progressing challenges [3, 4].

Clustering and auto-encoders are very important for anomaly detection in cybersecurity to avoid numerous threats. Cyber attacks can be identified by recognizing unusual system behavior. Traditional strategies such as Signature-Based Detection (SBD) are used to avoid previously known passive attack patterns but leave systems helpless to new and advanced active transaction threats. Anomaly discovery methods, especially those utilizing ML, offer a more vital tool to distinguish deviations from typical behavior, indicating malicious activities [5]. Deep Learning (DL)-based models have essentially increased the accuracy of anomaly detection. DL models, such as Convolutional Neural Networks (CNNs), autoencoders, and Generative Adversarial Networks (GANs), have appeared to perform satisfactorily in different areas [6]. Convolutional Neural Networks (CNNs) are known to extract features at various layers and effectively detect irregularities in complex and high-dimensional datasets. Deep spatial autoencoding models can be used efficiently to capture data variability [7]. Nonlinear autoencoders consist of an encoder that compresses the data into a lower-dimensional latent space and a decoder that recreates the input from this inactive representation and avoids irregularities [8, 9]. Variational autoencoders and gradient descent along with CNNs can be used to reduce data dimensionality and identify anomalies [9, 10].

II. LITERATURE REVIEW

Convolutional layers are used for anomaly detection usually based on an encoder and a decoder. CNN-based autoencoders are helpful in image processing and the collection of spatial information. Ensemble-based ML methods increase the model's efficacy and capacity to recognize inconsistencies in complex datasets [11, 12]. High-dimensional datasets require preprocessing steps, such as overseeing missing values, normalization, etc. These preprocessing steps are fundamental for the model's execution and information quality [3, 13]. In a CNN autoencoder, convolutional layers are used for feature extraction. Hyperparameters are used to optimize this procedure and ensure the detection of outliers or inconsistencies in the data [4]. CNN autoencoders can achieve improved accuracy, precision, and F1 score. These metrics comprehensively assess the model's ability to recognize quality and irregularities in various datasets. This study focuses on a CNN-based autoencoder for anomaly detection, tested and verified using a standard benchmark dataset.

III. METHOD

A. Dataset Description and Preprocessing

This study used the PTB Suggestive ECG Dataset [14], which comes from Physionet and is openly available. 14,552 samples were used, divided into two classes: those with ordinary pulses (Normal) and those with cardiovascular varieties (Anomalous). The full dataset consists of 21,837 samples, but using it would require more processing power and dividing it into three classes, which is left for future research. The ECG signals are assessed at 100 Hz, giving significant standard information that is reasonable for point-by-point assessment and model readiness. Table I shows the description of the dataset. Data preprocessing included standardization by segregating the data information into various testing sets.

TABLE I. DATASET DESCRIPTION

Feature	Description
Number of samples	14,552
Number of categories	2 (Normal, Anomalous)
Sampling frequency	100 Hz
Data source	Physionet's PTB diagnostic database

B. Proposed Autoencoder Convolutional Neural Network (AECNN) Architecture

An efficient AECNN architecture was proposed to detect anomalies, including various convolutional layers for feature extraction, clustering layers for data compression, and upsampling layers for data reconstruction. The encoder compresses the input data into a lower-dimensional representation, capturing essential features while discarding noise. The decoder reconstructs them from this representation, with the reconstruction error used to identify anomalies.

$$Reconstruction\ Error = \| \mathrm{X} - \widehat{\mathrm{X}} \|^2 \qquad (1)$$

where X is the original input and $\widehat{\mathrm{X}}$ is the reconstructed output.

1) Encoder

- Input layer: Accepts the ECG signal data with a shape of (None, 125).
- Convolutional layer 1: 32 filters, kernel size 3, activation function ReLU.
- MaxPooling layer 1: Pool size 2.
- Convolutional layer 2: 64 filters, kernel size 3, activation function ReLU.
- MaxPooling layer 2: Pool size 2.

2) Decoder

- Dense layer: 125 neurons, ReLU activation function, conforms the output to the dimensions of the input.
- UpSampling layer 1: Upsamples the information to match the result size of the principal MaxPooling layer.
- Convolutional Transpose Layer 1: 64 channels, size 3, ReLU activation function.
- UpSampling Layer 2: Upsamples the information to match the result size of the second MaxPooling layer.
- Convolutional Transpose Layer 2: 32 channels, size 3, ReLU activation function. Hyperparameter tuning was used to optimize its performance by minimizing the loss function

$$\text{Loss Function } = \frac{1}{N}\sum_{i=1}^{N}\left(X_i - \hat{X}_i\right)^2 \tag{2}$$

where N is the number of samples.

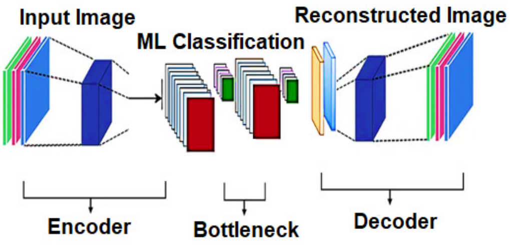

Fig. 1. Proposed Autoencoder Convolutional Neural Network (AECNN).

C. Evaluation Metrics

The effectiveness of the CNN autoencoder model was evaluated using various metrics, including accuracy, precision, recall, F1-score, and Mean Squared Error (MSE).

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \tag{3}$$

$$Precision = \frac{TP}{TP+FP} \tag{4}$$

$$Recall = \frac{TP}{TP+FN} \tag{5}$$

$$F1-score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \tag{6}$$

IV. EXPERIMENTAL RESULTS AND DISCUSSION

The proposed model was proficient in differentiating the two classes. Its AUC score [15] was close to 1, showing that it achieves a high True Positive Rate (TPR) and a low False Positive Rate (FPR), ensuring that irregularities and issues in the dataset can be easily rectified using the proposed autoencoder. The confusion matrix indicates that the model had a high TPR and a moderately low FNR, demonstrating compelling inconsistency recognition capacities. 7604 (72.38%) of TP anomalies were correctly identified, but there were 267 FN (6.60%). This can be confirmed by the confusion matrix shown in Figure 2.

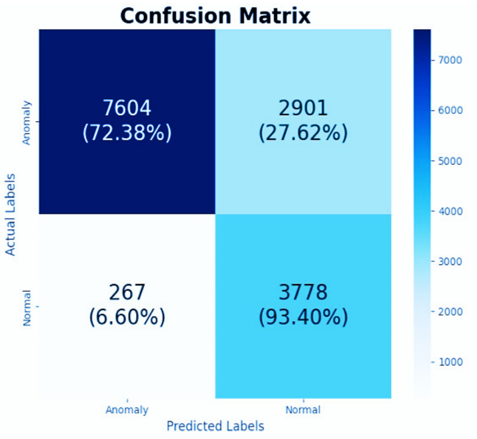

Fig. 2. Confusion matrix for AECNN.

A. Distribution of Reconstruction Errors

The reconstruction error plot provides a reasonable perception of how the CNN autoencoder model recognizes ordinary and irregular information. The cut-off point for identifying anomalous instances is 0.020, which is the threshold for anomaly detection. For the preparation of typical information, there is a high centralization of low recreation errors, showing that the model successfully learns and recreates ordinary examples during preparation. The Receiver Operating Characteristic (ROC) curve, which plots the TPR against the FPR at different limit settings, is a graphical representation of a model's detection ability.

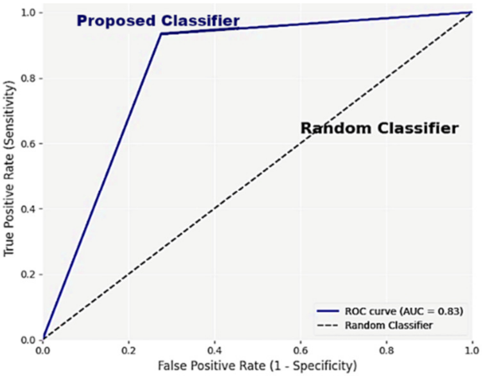

Fig. 3. ROC curve.

The capacity of the CNN autoencoder to distinguish between typical and atypical data is demonstrated by its AUC of 0.83. Figure 3 shows the ROC curve for a random classifier and the one proposed. The model's strength and steadfastness in the detection of real-world inconsistencies are reflected within the AUC of 0.83.

Figure 4 shows the reconstruction error distribution, which is a critical metric to assess an autoencoder's performance in anomaly detection, showing three different curves: test normal, train normal, and anomaly. High reconstruction errors typically indicate anomalies. Anomalies can be detected with a high degree of accuracy by evaluating the reconstruction error. This differentiation is essential for the model's reliability in practical applications. The fact that the test normal data have a similar distribution indicates that the model can be applied to unseen normal data with ease and has a low reconstruction error. Figures 5 and 6 show that the CNN autoencoder model is trained on data by improving its presentation and keeping up with its adequacy when applied to irregular data. The results show normal losses of 0.010, 0.015, 0.013, 0.011, and 0.185.

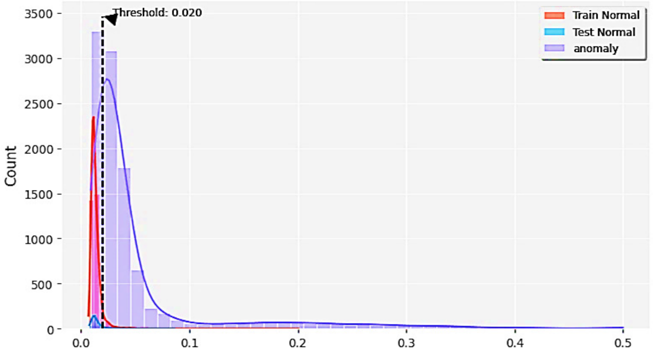

Fig. 4. Reconstruction error distribution.

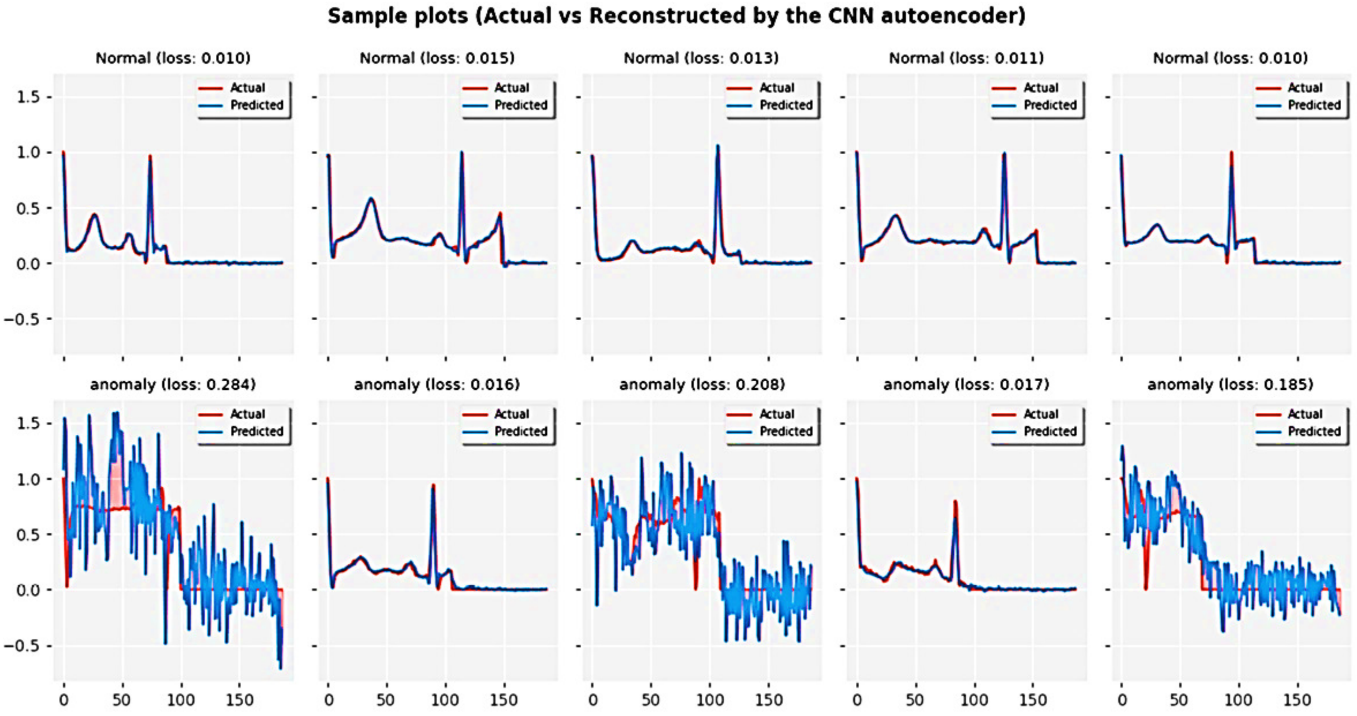

Fig. 5. Sample plots: actual vs. reconstructed by the proposed AECNN.

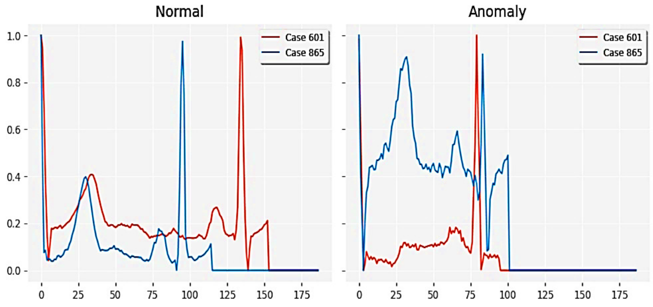

Fig. 6. Plots for normal vs anomaly cases.

Figure 6 shows two plots, where the blue line represents the normal peak with low errors for Normal data and the red line represents the Anomalous data. The two data instances are examined from the benchmark dataset, showing low errors for Normal and high recreation errors for Anomalous. This examination highlights the model's precision in reproducing typical examples and its battle against abnormalities, actually recognizing the two because of recreation errors. Two dedicated cases were evaluated and tested on the AECNN model, showing reasonable results. In normal case 601, the reconstruction error was outstandingly low, demonstrating that the model precisely recreated the data. Alternately, in the anomalous case 865, the recreation error was essentially higher. This significant error suggests that there is an anomaly and that the model is having trouble accurately reconstructing the input data.

Figure 7 presents the training and validation loss. The x-axis represents the epochs, ranging from 0-70, while the y-axis represents the loss of Anomalous data that is recognized by the reconstruction examination. Reconstruction errors for typical data are consistently low, indicating that the model accurately captures and imitates typical examples. The model's productive limit for learning and reconstruction is visible and consistent.

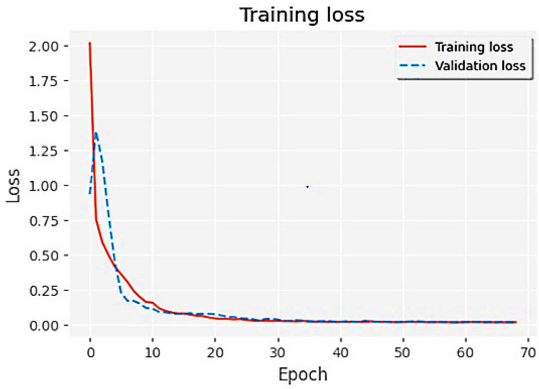

Fig. 7. Training and validation loss.

As shown in Figure 7, the training and validation processes show the optimization of the model through various epochs and adjusting hyperparameters to minimize the reconstruction error as shown in (7).

$$Training\ Loss = \frac{1}{N}\sum_{i=1}^{N} \|\mathrm{X}_i - \widehat{\mathrm{X}}_i\|^2 \qquad (7)$$

where N is the number of samples, X_i is the original input, and $\hat{X}_i$ is the reconstructed output. In addition, the Adam optimizer was used for training with a learning rate of 0.001.

$$\theta_{t+1} = \theta_t - \eta\nabla_\theta J(\theta) \qquad (8)$$

where θ represents the model parameters, η is the learning rate, and $J(\theta)$ is the loss function.

Figure 8 shows a smoothed mean plot for each class, both Normal and Anomalous. The model's ability to recognize anomalies is exhibited by its expanded changeability and reduced errors. Figure 9 presents a histogram, indicating a well-defined peak in the lower range that suggests a successful reconstruction of Normal and Anomalous instances.

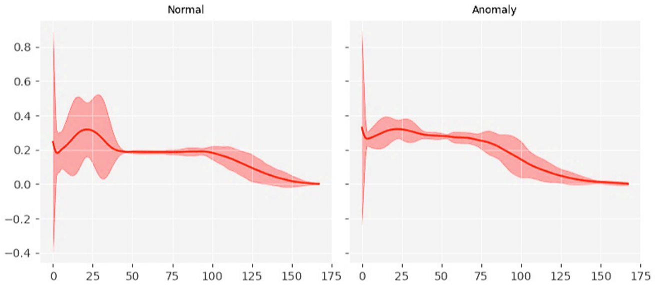

Fig. 8. Smoothed mean plots for each class.

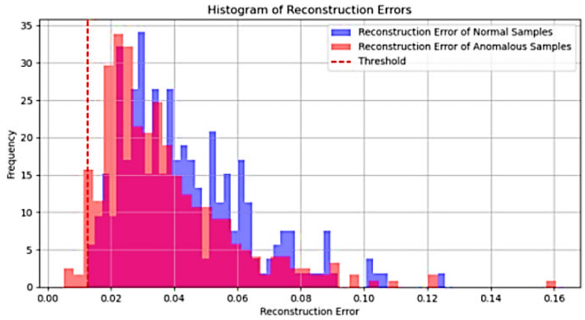

Fig. 9. Histogram of reconstruction errors with a threshold of 0.012.

B. *Evaluation Matrix*

1) *Reconstruction Error*

Reconstruction error was calculated as

$$E = \|X - X'\|^2 = \| [1.0, 0.9, \dots] - [0.98, 0.87, \dots] \|^2 \approx 0.0025$$

2) *Loss Function*

The training loss function calculates the loss for the number of samples $N = 14552$

$$Loss = \frac{1}{14552}\sum_{i=1}^{14552} (X_i - X_i')^2 \approx 0.015$$

3) *Evaluation Metrics*

The evaluation metrics were calculated based on the values shown in the confusion matrix (Figure 2). Table II presents the evaluation metrics for the proposed model. These metrics show the model's enhancement in terms of accuracy by precisely recognizing anomalies while keeping a low pace FP and FN.

$$\text{Accuracy} = \frac{7604+3778}{7604+3778+2901+267} \approx 0.7821$$

$$\text{Precision} = \frac{7604}{7604+2901} \approx 0.7238$$

$$\text{Recall} = \frac{7604}{7604+267} \approx 0.9641$$

$$\text{F1-Score} = \frac{2\cdot 0.7238 - 0.9641}{0.7238+0.9641} \approx 0.8261$$

Given the provided ROC curve data, the AUC is approximately:

$$\text{AUC} \approx 0.83$$

TABLE II. EVALUATION METRICS

Metric	Value	Metric	Value
Accuracy	78.21%	Recall (TPR)	96.40%
Precision	72.38%	F1-Score	82.61%

V. CONCLUSION

This study demonstrated the effectiveness of using CNN-based autoencoders for anomaly detection in high-dimensional datasets, especially on an available electrocardiography dataset. The model combines the qualities of convolutional layers to extract features for unsupervised learning. The experimental results, highlighted by accuracy, precision, recall, F1-score, and a critical AUC of 0.83, affirm the model's ability to distinguish anomalies with high reconstruction errors. The design of the proposed CNN autoencoder and data preprocessing were essential for robust performance. The ability to capture perplexing spatial connections inside the information made the CNN autoencoder exceptionally effective, which was tested employing a dataset that presents real-world complexities. The model's low FN rate and solid execution measurements emphasize its potential for applications in fields requiring solid anomaly detection, such as restorative diagnostics, finance, and cybersecurity. Future work should focus on improving the model's architecture, exploring diverse autoencoder types, and applying the proposed approach to other high-dimensional datasets to advance its strength and extend its applicability.

REFERENCES

[1] M. I. H. Okfie and S. Mishra, "Anomaly Detection in IIoT Transactions using Machine Learning: A Lightweight Blockchain-based Approach," *Engineering, Technology & Applied Science Research*, vol. 14, no. 3, pp. 14645–14653, Jun. 2024, https://doi.org/10.48084/etasr.7384.

[2] P. More and P. Mishra, "Enhanced-PCA based Dimensionality Reduction and Feature Selection for Real-Time Network Threat Detection," *Engineering, Technology & Applied Science Research*, vol. 10, no. 5, pp. 6270–6275, Oct. 2020, https://doi.org/10.48084/etasr.3801.

[3] V. Chandola, A. Banerjee, and V. Kumar, "Anomaly detection: A survey," *ACM Computing Surveys*, vol. 41, no. 3, Apr. 2009, Art. no. 15, https://doi.org/10.1145/1541880.1541882.

[4] V. Hodge and J. Austin, "A Survey of Outlier Detection Methodologies," *Artificial Intelligence Review*, vol. 22, no. 2, pp. 85–126, Oct. 2004, https://doi.org/10.1023/B:AIRE.0000045502.10941.a9.

[5] M. H. Bhuyan, D. K. Bhattacharyya, and J. K. Kalita, "Network Anomaly Detection: Methods, Systems and Tools," *IEEE Communications Surveys & Tutorials*, vol. 16, no. 1, pp. 303–336, 2014, https://doi.org/10.1109/SURV.2013.052213.00046.

[6] C. Zhou and R. C. Paffenroth, "Anomaly Detection with Robust Deep Autoencoders," in *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, Halifax, Canada, Aug. 2017, pp. 665–674, https://doi.org/10.1145/3097983.3098052.

[7] C. Baur, B. Wiestler, S. Albarqouni, and N. Navab, "Deep Autoencoding Models for Unsupervised Anomaly Segmentation in Brain MR Images," in *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries*, Granada, Spain, 2019, pp. 161–169, https://doi.org/10.1007/978-3-030-11723-8_16.

[8] M. Sakurada and T. Yairi, "Anomaly Detection Using Autoencoders with Nonlinear Dimensionality Reduction," in *Proceedings of the MLSDA 2014 2nd Workshop on Machine Learning for Sensory Data Analysis*, Gold Coast, Australia, Dec. 2014, pp. 4–11, https://doi.org/10.1145/2689746.2689747.

[9] J. An and S. Cho, "Variational autoencoder based anomaly detection using reconstruction probability," SNU Data Mining Center, Special Lecture on IE, 2015.

[10] G. E. Hinton and R. R. Salakhutdinov, "Reducing the Dimensionality of Data with Neural Networks," *Science*, vol. 313, no. 5786, pp. 504–507, Jul. 2006, https://doi.org/10.1126/science.1127647.

[11] M. H. H. Khairi, S. H. S. Ariffin, N. M. A. Latiff, A. S. Abdullah, and M. K. Hassan, "A Review of Anomaly Detection Techniques and Distributed Denial of Service (DDoS) on Software Defined Network (SDN)," *Engineering, Technology & Applied Science Research*, vol. 8, no. 2, pp. 2724–2730, Apr. 2018, https://doi.org/10.48084/etasr.1840.

[12] J. Masci, U. Meier, D. Cireşan, and J. Schmidhuber, "Stacked Convolutional Auto-Encoders for Hierarchical Feature Extraction," in *Artificial Neural Networks and Machine Learning – ICANN 2011*, Espoo, Finland, 2011, pp. 52–59, https://doi.org/10.1007/978-3-642-21735-7_7.

[13] U. Khan, K. Khan, F. Hassan, A. Siddiqui, and M. Afaq, "Towards Achieving Machine Comprehension Using Deep Learning on Non-GPU Machines," *Engineering, Technology & Applied Science Research*, vol. 9, no. 4, pp. 4423–4427, Aug. 2019, https://doi.org/10.48084/etasr.2734.

[14] P. Wagner, N. Strodthoff, R.-D. Bousseljot, W. Samek, and T. Schaeffter, "PTB-XL, a large publicly available electrocardiography dataset." PhysioNet, https://doi.org/10.13026/X4TD-X982.

[15] D. M. Powers, "Evaluation: From precision, recall and F-measure to ROC, informedness, markedness and correlation," *Journal of Machine Learning Technologies*, vol. 1, no. 1, pp. 37–63, 2011.