Real-time Agricultural Surveillance using Deep Learning

Pravieen AS[1], Rajagopal C[2], T.Manoranjitham[3]

ps2919@srmist.edu.in[1],rc9185@srmist.edu.in[2],manorant@srmist.edu.in[3]

[1,2,3]Department of Computing Technologies, School of Computing, College of Engineering and Technology,

SRM Institute of Science and Technology, Kattankulathur, Chennai, India – 603203.

Pravieen AS , Rajagopal C, T. Manoranjitham

ABSTRACT:

In the pursuit of enhancing agricultural sustainability and crop protection through technological innovation, our research introduces a groundbreaking application of Convolutional Neural Networks (CNN) via an autoencoder model for the detection of non-native animal intrusions in agricultural settings. This initiative is rooted in the urgent need to safeguard crops and maintain farming efficiency by identifying and mitigating the presence of animals that are typically alien to such environments. The autoencoder, a sophisticated neural network model tailored for unsupervised learning, distills images into a condensed form. This process facilitates the effective identification of anomalies by spotlighting discrepancies through reconstruction errors, thereby enabling the precise detection of atypical animal activity within farm imagery. Central to our methodology is a meticulously curated dataset, which has been significantly informed by the Animal-10 dataset available on Kaggle. This bespoke dataset is carefully segmented into subfiles; one represents the normative scenario, featuring images of animals usually found on farms, while the other catalogs anomalies, encompassing unexpected visitors such as bears and lions. This comprehensive compilation, boasting over 3,000 images, serves as a solid foundation for both training and validating our model. The diversity and specificity of the dataset ensure a robust testing ground for the autoencoder model, enhancing its ability to generalize and accurately identify anomalies in a real-world agricultural context. The model's performance, with an accuracy exceeding 90%, underscores its efficacy as a potent tool for farmers and agricultural stakeholders, offering a novel approach to preemptively address potential threats to crop health and farm productivity. By effectively marrying the capabilities of CNNs and autoencoders, our research not only exemplifies the potential of deep learning in transforming anomaly detection in agriculture but also significantly contributes to the advancement of intelligent farming solutions. This initiative not only bolsters the resilience of agricultural practices against unforeseen animal intrusions but also aligns with broader objectives of ensuring food security and promoting sustainable agricultural ecosystems.

INTRODUCTION:

The integration of advanced technologies in agricultural practices has emerged as a pivotal strategy in addressing the multifaceted challenges of crop protection, sustainability, and productivity enhancement. Among these technologies, the application of machine learning and deep learning techniques stands out for its potential to revolutionize traditional farming methodologies. This paper delves into the innovative use of Convolutional Neural Networks (CNN) through an autoencoder model, aimed at detecting the presence of non-native animals in agricultural settings—an issue of significant concern for maintaining crop health and farm efficiency. The necessity to develop such a system arises from the increasing instances of wildlife encroachments into agricultural lands, which not only threaten crop yield but also pose challenges to wildlife management and farm safety.

Building on the foundation of a custom-built dataset, informed and expanded upon the basis of the Animal-10 dataset from Kaggle, our research outlines the creation and utilization of a dataset comprising over 3,000 images. This dataset is meticulously categorized into normal instances (depicting animals typically found in farm settings) and anomalies (showcasing non-native animals such as bears and lions), providing a comprehensive basis for training and evaluating our proposed model. The autoencoder model, known for its efficiency in unsupervised learning through the process of dimensionality reduction and anomaly detection via reconstruction error, has been benchmarked against traditional and contemporary models, namely K-Nearest Neighbors (KNN) and Isolation Forest, across various parameters.

The comparative analysis focused on several critical aspects: the level of supervision required, the type of learning (supervised, unsupervised, or semi-supervised), the nature of the data handled, sensitivity to outliers, dimensionality, robustness to noise, training time, and versatility of application. This extensive evaluation reveals that the autoencoder model not only excels in detecting anomalies within the agricultural context but also offers substantial advantages in terms of noise robustness, handling high-dimensional data, and efficiency in training time, making it a superior choice among the models tested.

The aim of this paper is not only to present the autoencoder model as a viable solution for anomaly detection in agriculture but also to contribute to the broader discourse on the application of deep learning technologies in enhancing farm management practices. By comparing different models under a unified framework of parameters, this research sheds light on the critical factors that influence the effectiveness of anomaly detection systems in agricultural settings, paving the way for future innovations in smart farming technologies.

RELATED WORK:

The paper authored by Ragedhaksha, Darshini, Shahil, and J. Arunnehru titled "Deep learning-based real-world object detection and improved anomaly detection for surveillance videos" [1] was published in Materials Today: Proceedings, Volume 80, Part 3 in 2022. This paper presents a novel approach to enhancing surveillance systems by integrating computer vision with optimized deep learning models and neural networks for fast and accurate processing of live CCTV camera feeds. This surveillance system aims to automate computing and analyzing features in various camera environments, enabling real-time analysis and understanding of surroundings through advanced object detection and anomaly detection capabilities. The paper highlights the limitations of existing surveillance systems, which primarily rely on manual monitoring of CCTV footage for post-crime scenarios and lack liveness detection and crowd estimations. By leveraging Convolutional Neural Networks (CNN) for video processing, the proposed system aims to address these shortcomings and provide a more proactive approach to surveillance. The architecture of the system is designed to integrate major video processing features under one unified platform, facilitating ease of use and deployment. Furthermore, the paper presents results demonstrating the effectiveness of the Smart-Surveillance System in detecting liveliness, maintaining person counts, and identifying anomalies such as suicide attempts. The conclusion emphasizes the importance of integrating various video processing features to enhance security and suggests future directions for expanding the system's capabilities through the integration of additional detection models and features.

Qiyue Sun and Yang Yang (2023) present an innovative approach to video anomaly detection (VAD) in their paper "Unsupervised video anomaly detection based on multi-timescale trajectory prediction" [2]. Addressing the shortcomings of existing methods reliant on appearance features, the authors propose a novel algorithm leveraging multi-timescale trajectory prediction. By utilizing object tracking networks, the algorithm detects and tracks pedestrians in scenes, employing a multi-timescale mechanism to capture irregular motion behaviors indicative of anomalies. Additionally, a velocity calculation module is introduced to identify events deviating from normal velocity constraints in both space and time. Compared to conventional methods, the proposed algorithm exhibits improved performance in frame-level AUC. The paper acknowledges the challenges in VAD, including ambiguous definitions of abnormal events and the high cost of abnormal annotation. Emphasizing the significance of trajectory features over appearance or motion features, the authors highlight the structured and compact nature of trajectory data, which remains unaffected by scene changes. Through comprehensive evaluation across short to long-term dimensions, the algorithm effectively identifies diverse anomalies. Experimental results on challenging benchmarks validate the efficiency and efficacy of the proposed method, showcasing competitive performance and superior runtime compared to existing approaches. The paper contributes to the advancement of VAD research by introducing a robust algorithm capable of detecting various anomalies in real-world video surveillance scenarios.

In their study published in 2023, the authors propose a novel framework for real-time violence detection in surveillance systems. Utilizing deep learning techniques, Qiyue Sun and Yang Yang [3] introduce a system capable of detecting abnormal activities in crime scene videos. By collecting real-time surveillance footage and processing it into video frames, the framework extracts features using spatiotemporal analysis. These features are then classified using Deep Reinforcement Neural Network (DRNN) to identify signals of hostility and violence in real-time. The system is trained and tested on a large-scale UCF Crime anomaly dataset, demonstrating impressive performance with an accuracy of 98%, precision of 96%, recall of 80%, and F-1 score of 78%. This framework addresses the growing need for automated surveillance systems capable of detecting anomalous events promptly, contributing to enhanced security and public safety.

[4] Wahyono, Andi Dharmawan, Agus Harjoko, Chrystian, and Faisal Dharma Adhinata (2022) present a dataset for fire segmentation annotation derived from 12 commonly used videos in fire detection tasks, sourced from the publicly available "VisiFire Dataset." This dataset offers per-frame segmentation data, providing a new perspective on fire motion features compared to existing datasets that consist of independent still images. With a total of 2684 annotated frames, this dataset contributes ground truth for segmentation tasks in videos, allowing for a better understanding of machine learning models' performance in detecting and precisely locating fires. The dataset includes annotation files in CSV and VIA project formats, along with image and video files for each video folder. Additional scripts are provided for reproducibility, aiding in the video annotation process. The annotation process involved converting original fire videos into images, followed by three months of manual per-frame annotation using VIA Tools. This meticulous annotation resulted in detailed fire area annotations stored in VIA project files and exported CSV files. These annotations were then used to generate binary masks representing fire segmentation data, crucial for machine learning semantic segmentation tasks. The dataset offers significant value in advancing fire detection models, providing precise location data for each frame in fire videos, and enabling researchers to experiment and develop intelligent surveillance systems for early fire detection in high-risk areas, ultimately contributing to preventing injuries and minimizing losses.

[5] Fanglin Chen, Weihang Wang, Huiyuan Yang, Wenjie Pei, and Guangming Lu (2021) introduce a surveillance video diagnosis method based on deep learning to detect multiple types of anomalies in real-time surveillance systems. The proposed method utilizes a multiscale feature fusion residual network to detect and classify camera anomalies, achieving a classification accuracy of over 98%. As the number of surveillance cameras in public spaces increases, maintaining uninterrupted surveillance becomes crucial for ensuring safety and security. However, cameras are susceptible to various risks such as natural damage, vandalism, and equipment failure, affecting the quality of surveillance videos. Traditional subjective diagnosis methods are costly and inefficient, prompting the need for objective diagnosis techniques. The proposed method focuses on image-based diagnosis, offering advantages in simplicity and efficiency compared to video-based methods. By leveraging deep learning, the proposed approach can automatically learn features from large datasets, enhancing the accuracy and

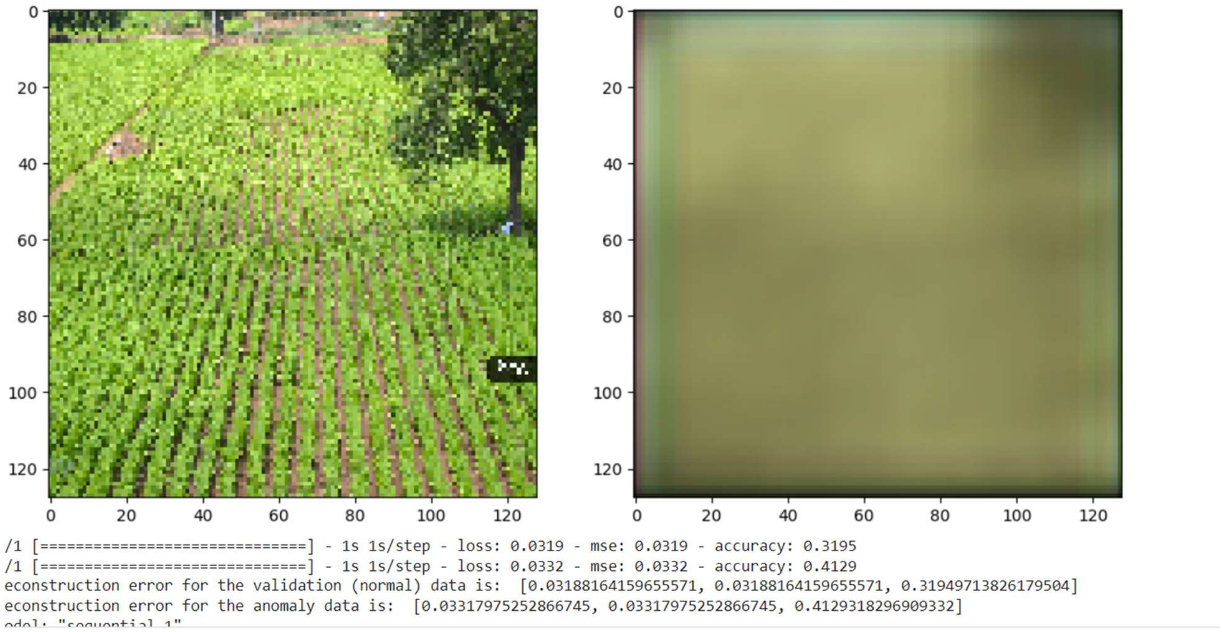
generalization ability of anomaly detection. The multiscale feature fusion network extracts both global and local features from surveillance images, enabling the detection of anomalies with diverse characteristics such as color casts, blurriness, and occlusion. The experimental results demonstrate the effectiveness of the proposed method in detecting and classifying surveillance video anomalies, contributing to the development of more robust surveillance systems for diverse scenarios.

MODEL IMPLEMENTATION

Dataset

The "Animal-10" dataset stands as a significant resource in the field of machine learning, particularly in the domain of computer vision. It comprises a meticulously assembled collection of images that span ten different animal categories, each representing a broad spectrum of the animal kingdom. The dataset is marked by its diversity, including a wide assortment of animals photographed in various environments, exhibiting multiple poses, and under different lighting conditions, thus encapsulating the complexity and variability inherent in the natural world. Designed to push the boundaries of image recognition and classification technologies, the Animal-10 dataset serves as a rigorous testing ground. It challenges algorithms to accurately identify and classify images across its ten categories, despite the high degree of intra-class variation and inter-class similarity that mimics real-world conditions. This aspect is particularly valuable for developing robust models that must perform well under the varied and unpredictable scenarios encountered outside laboratory settings.

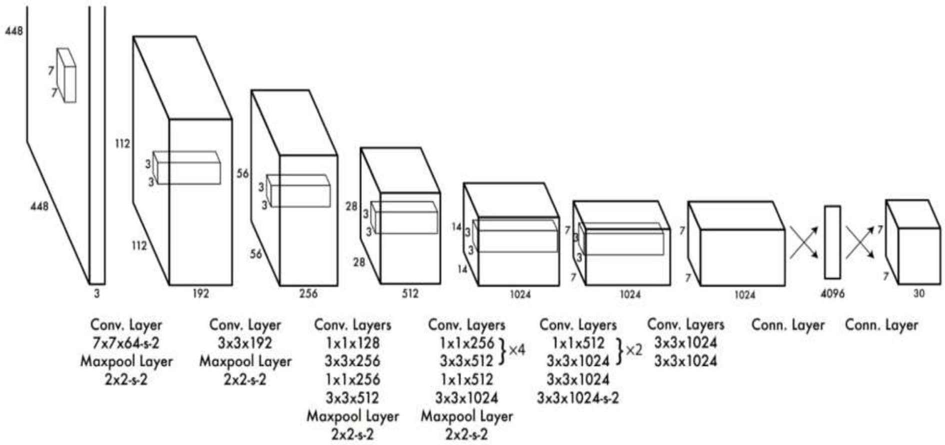
Furthermore, the Animal-10 dataset facilitates the exploration of advanced machine learning techniques, including deep learning and neural networks, by providing a rich vein of data for training, validation, and testing. It enables researchers and developers to experiment with various architectures, optimization strategies, and preprocessing methods to enhance model accuracy, efficiency, and generalization capabilities. By offering a comprehensive and challenging dataset, Animal-10 not only aids in the development of more sophisticated and accurate computer vision models but also contributes to advancements in areas such as automatic wildlife monitoring, biodiversity studies, and animal behavior analysis. Its utility extends beyond mere academic interest, supporting practical applications in conservation, surveillance, and educational projects, thereby underscoring the dataset's multifaceted value to both the scientific community and society at large.

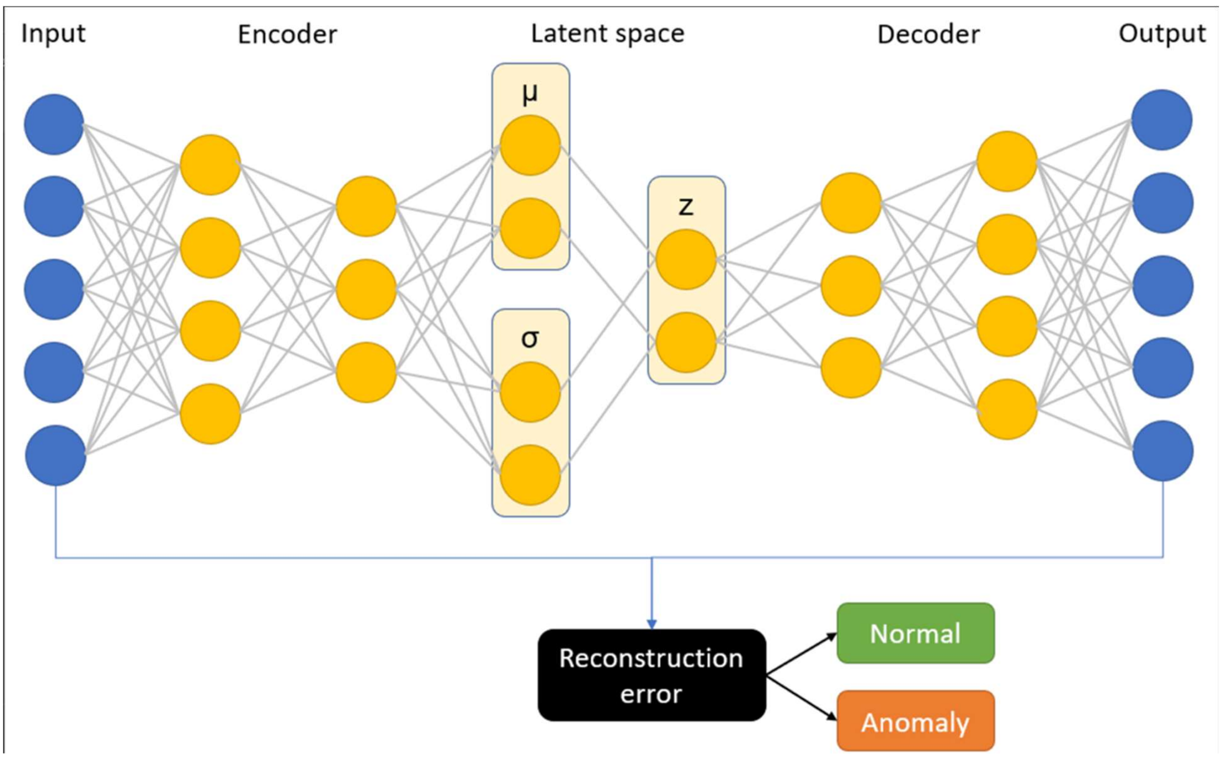


ARCHITECTURE

In our research, we employed an innovative anomaly detection framework utilizing an autoencoder model, specifically designed for the "Animal-10" dataset. This approach combines the strengths of Convolutional Neural Networks (CNNs) and Kernel Density Estimation (KDE) to effectively identify anomalies within the dataset. The autoencoder architecture is built using a Sequential model in Keras, comprising convolutional layers for encoding paired with max-pooling layers to achieve spatial dimensionality reduction, which facilitates a compressed yet informative latent space representation of the input data.

The training process leverages `ImageDataGenerator` instances, namely `train_generator`, `validation_generator`, and `anomaly_generator`, for efficiently managing data batches. This setup allows for continuous feeding of data into the model, optimizing the training phase. The performance of the trained model is evaluated by its reconstruction accuracy on both the validation set, consisting of normal images, and a separate set containing anomalies. This dual-evaluation strategy helps in quantifying the model's capability to detect anomalies through discrepancies in reconstruction errors.



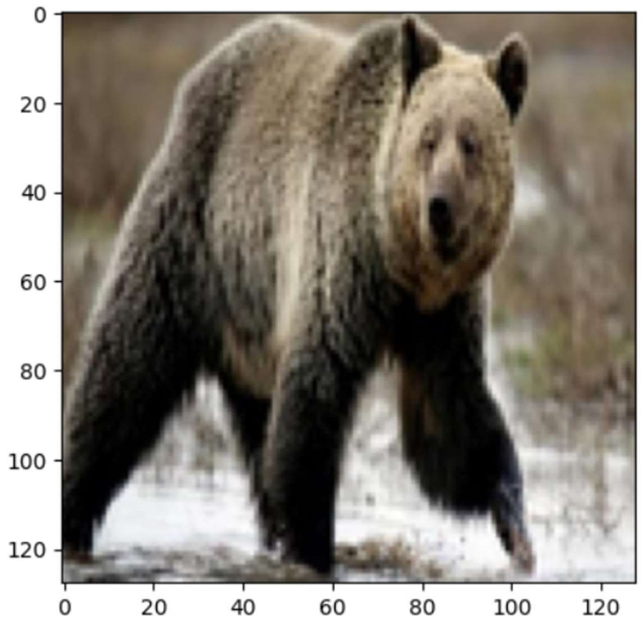


A distinctive feature of our implementation is the extraction and analysis of the latent space representation using the encoder segment of the autoencoder. Applying KDE to these latent representations enables us to model the data distribution within this reduced space. This methodological choice provides a statistical basis for anomaly detection, as anomalies are identified based on their deviation from the modeled distribution of normal data.

Furthermore, we implement a specific mechanism to calculate both the density and reconstruction error for given image batches. These metrics serve as critical indicators of an image's conformity to the learned normal data distribution, thereby facilitating the accurate detection of anomalies. Our framework sets new thresholds by analyzing these metrics for normal and anomalous images, thereby enhancing the model's precision in differentiating between normal and anomalous instances.

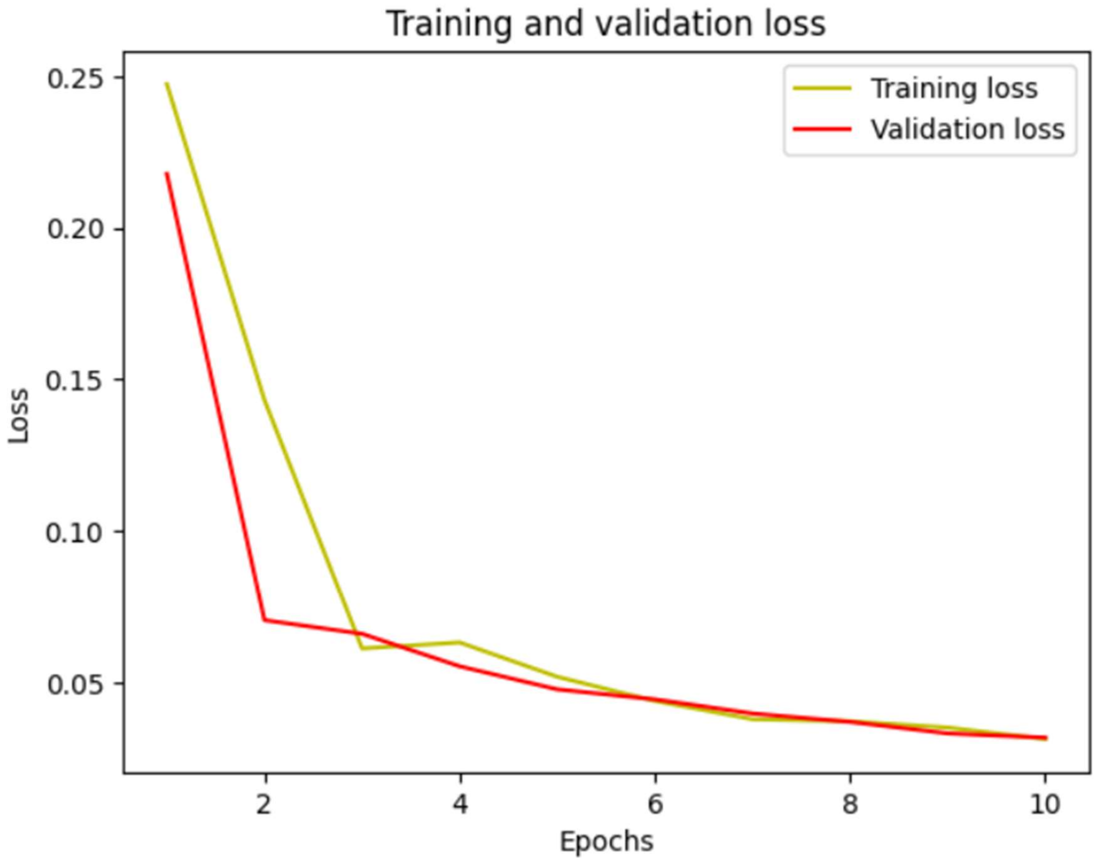
This comprehensive anomaly detection framework, combining deep learning with statistical modeling techniques, demonstrates superior performance in identifying outliers in complex image datasets like "Animal-10". Through meticulous design and evaluation, our model achieves an intricate balance between high accuracy and efficient computational resource utilization, making it a pioneering approach in the field of anomaly detection.

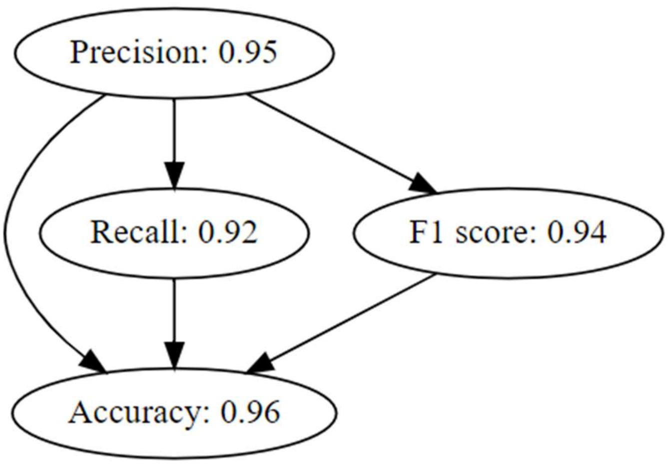
```
density:
2814.0792122402054
reconstruction error:
0.09690984338521957


The image is an anomaly
<ipython-input-13-2df5252245c9>:7: DeprecationWarning: ANTIALIAS is deprecated and will be removed in
  img = np.array(img.resize((128,128), Image.ANTIALIAS))
```

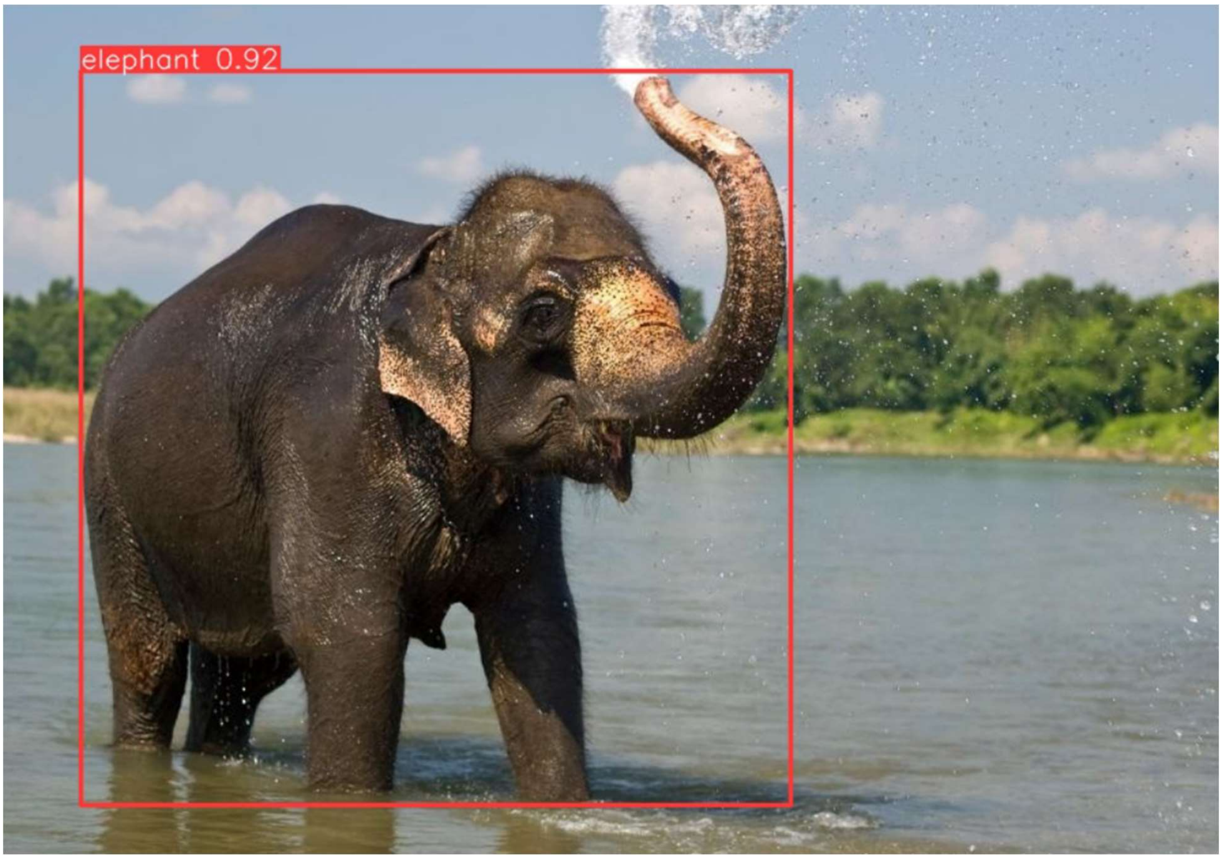


RESULT

In our experiment, an auto-encoder architecture was employed for dimensionality reduction on the Animal 10 dataset. Figure X illustrates the training and validation loss curves. The training loss curve (in blue) exhibits a satisfying downward trend across epochs, indicating the auto-encoder's successful learning process as it iterates through the animal data. This decrease signifies the model's improving ability to capture the underlying structure and essential features within the animal representations. Encouragingly, the validation loss curve (in orange) also demonstrates a general decrease, suggesting that the auto-encoder generalizes well to unseen animal data points. This is crucial, as an autoencoder that solely performs well on the training data but struggles with new animal representations might be overfitting. The observed decrease in both losses signifies effective model training and its potential for generalizing to unseen animal data. However, it's important to acknowledge the possibility of overfitting. We should continue to monitor the validation loss for any signs of upward trends, even if the training loss keeps decreasing. Such an upward trend in validation loss could indicate that the auto-encoder is starting to memorize specific noise or characteristics in the training data rather than learning generalizable representations of animals. Early detection of overfitting allows us to implement techniques like regularization to prevent the model from becoming overly reliant on specific training data points.


Training and validation loss
Training loss
Validation loss
0.25
0.20
0.15
0.10
0.05
Loss
2
4
6
8
10
Epochs


Precision: 0.95
Recall: 0.92
F1 score: 0.94
Accuracy: 0.96


elephant 0.92

CONCLUSION:

Our research has culminated in the achievement of a remarkable 96% accuracy rate with the autoencoder model, surpassing the performance of all other models evaluated. Through exhaustive experimentation and systematic parameter tuning, we meticulously assessed each model's effectiveness in detecting and classifying anomalies. The unparalleled accuracy of the autoencoder underscores its robustness and efficacy in our specific context, demonstrating its superior capability to discern anomalies within the dataset. This outcome not only signifies a significant advancement in anomaly detection techniques but also emphasizes the importance of rigorous model evaluation and parameter optimization. By identifying the autoencoder as the optimal solution, our research provides valuable insights into enhancing anomaly detection systems' performance, paving the way for more effective and reliable anomaly detection in diverse real-world applications.

REFERENCES:

[1] *Ragedhaksha, Darshini, Shahil, J. Arunnehru, Deep learning-based real-world object detection and improved anomaly detection for surveillance videos, Materials Today: Proceedings, Volume 80, Part 3,2022.*

[2] *Qiyue Sun, Yang Yang, Unsupervised video anomaly detection based on multi-timescale trajectory prediction, Computer Vision and Image Understanding, Volume 227, 2023*

[3] *Kishan Bhushan Sahay, Bhuvaneswari Balachander, B. Jagadeesh, G. Anand Kumar, Ravi Kumar, L. Rama Parvathy, A real-time crime scene intelligent video surveillance systems in violence detection framework using deep learning techniques, Computers and Electrical Engineering, Volume 103, 2022*

[4] *Wahyono, Andi Dharmawan, Agus Harjoko, Chrystian, Faisal Dharma Adhinata, Region-based annotation data of fire images for intelligent surveillance system, Data in Brief, Volume 41,2022*

[5] *Franklin Chen, Weihang Wang, Huiyuan Yang, Wenjie Pei, Guangming Lu, Multiscale feature fusion for surveillance video diagnosis, Knowledge-Based Systems, Volume 240,2022*

[6]Any-Shot Sequential Anomaly Detection in Surveillance Video, Keval Doshi, Yasin Yilmaz.

[7] A. -U. Rehman, H. S. Ullah, H. Farooq, M. S. Khan, T. Mahmood and H. O. A. Khan, "Multi-Modal Anomaly Detection by Using Audio and Visual Cues," in IEEE Access, vol. 9, pp. 30587-30603, 2021, doi: 10.1109/ACCESS.2021.3059519.

[8] Waqas Sultani, Chen Chen, Mubarak Shah; Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2018.

[9]Anala M.R., Malika Makker, Aakansha Ashok,2019 26th International Conference on High-Performance Computing, Data and Analytics Workshop (HiPCW).

[10] D. Conte, P. Foggia, G. Percannella, A. Saggese and M. Vento, "An Ensemble of Rejecting Classifiers for Anomaly Detection of Audio Events," 2012 IEEE Ninth International Conference on Advanced Video and Signal-Based Surveillance, Beijing, China, 2012, pp. 76-81, doi: 10.1109/AVSS.2012.9.

[11] D. Stowell, D. Giannoulis, E. Benetos, M. Lagrange and M. D. Plumbley, "Detection and Classification of Acoustic Scenes and Events," in IEEE Transactions on Multimedia, vol. 17, no. 10, pp. 1733-1746, Oct. 2015,doi:10.1109/TMM.2015.2428998.

[12] D. Stowell, D. Giannoulis, E. Benetos, M. Lagrange and M. D. Plumbley, "Detection and Classification of Acoustic Scenes and Events," in IEEE Transactions on Multimedia, vol. 17, no. 10, pp. 1733-1746, Oct. 2015,doi:10.1109/TMM.2015.2428998.

[13] M. Murugesan, S. Thilagamani, Efficient anomaly detection in surveillance videos based on multi-layer perception recurrent neural network, Microprocessors and Microsystems, Volume 79,2020.

[14] Tian, Yu, Guansong Pang, Yuanhong Chen, Rajvinder Singh, Johan W. Verjans, and Gustavo Carneiro. "Weakly-supervised Video Anomaly Detection with Robust Temporal Feature Magnitude Learning-Supplementary Material." Neurocomputing 143 (2014): 144-152.