Crime Detection with Variational Autoencoders

Pokkuluri Kiran Sree[1](✉), Majji Tejaswi[1], Ch Phaneendra Varma[1], Gurujukota Ramesh Babu[1], N SSSN Usha Devi[2], P. B. V. Raja Rao[1], M. Prasad[1], and P. J. R. Shalem Raju[1]

[1] Department of Computer Science and Engineering, Shri Vishnu Engineering College for Women, Bhimavaram, India
drkiransree@gmail.com
[2] Department of Computer Science and Engineering, University College of Engineering, JNTU, Kakinada, India

Abstract. The paper presents a novel idea of proposing the application of Variational Autoencoders (VAEs) in crime detection for predicting face aging and deaging, which is one of the potential challenge of forensic science. VAEs are equipped to generate realistic aged or rejuvenated facial images in simulating the appearance of suspects over time. Forensic age progression through VAEs could be pivotal in matching facial images with existing databases, potentially advancing criminal investigations. This work has considered learning rate and mean square error (MSE) which play a crucial role in determining the quality and efficiency of the image reconstruction process. Structural Similarity Index (SSI) and face recognition accuracy parameters are considered to evaluate the proposed work and to compare the work with the existing literature. The precision and recall of the proposed method is 0.923 & 0.936 which is promising when compared with the existing literature.

Keywords: Variational Autoencoders (VAEs) · Crimes · Image Processing

1 Introduction

Most of the researchers are using advanced machine learning techniques for detecting the crimes and preventing the same. Traditional mechanism depends on the analysis of crime data manually which may be prone to biases and consumes more time. After the arrival of generating models in deep learning such as Vibrational Autoencoders (VAEs) [1], there is big potential for more accurate and efficient crime detection systems. Artificial neural networks (ANNs) are one type VAE which are capable of learning to both encode and decode complex data distributions. A decoder network [2] reconstructs the input data from the latent space, while an encoder network maps input data to a latent space. It is feasible to teach a VAE useful representations of criminal activity and patterns by training it on crime data. The capacity of VAEs to produce realistic samples from the learnt data distribution is one of the main benefits of utilizing them for criminal

P. Pareek et al. (Eds.): IC4S 2024, LNICST 598, pp. 94–101, 2025.
https://doi.org/10.1007/978-3-031-77078-4_8

detection. This implies that a VAE [3] can produce synthetic crime data after it has been trained, which can be useful for enhancing the strength of crime detection models and augmenting small datasets.

Moreover, anomaly detection is a critical function for recognizing odd or suspicious activity that deviates from typical patterns is another feature that VAEs provide. Anomalies based on the reconstruction error can be found by comparing the reconstructed data with the original input [4]. In this study, we provide a unique VAE-based method for crime detection. We showcase how VAEs can outperform conventional crime detection algorithms and prove the efficacy of our approach on real-world crime datasets. We also go over the difficulties and practical aspects of implementing VAE-based crime detection systems in actual environments. All things considered, we think there is a lot of potential for increasing the effectiveness and precision of crime prevention initiatives when VAEs are incorporated into crime detection frameworks [5].

Many methods to identify the crimes are prevailing, statistical analysis is the process of looking for patterns, trends, and correlations in crime data by applying statistical techniques. To gain insights from crime data, methods including regression analysis, time series analysis, and spatial analysis are frequently employed. Crime detection tasks have been tackled by supervised learning algorithms such neural networks, support vector machines (SVM) [6], decision trees, and random forests. These algorithms categories or forecast criminal activity by using labelled crime data as a source of learning. This approach makes use of machine learning algorithms to predict the probable times and locations of crimes. In addition to other pertinent variables including demographics, socioeconomic indicators, and environmental characteristics, it frequently entails evaluating past crime data [7].

Variational autoencoders provides synthetic data for augmentation and anomaly detection in addition to learning rich representations of crime data in an unsupervised way. The effectiveness of crime prevention initiatives can be increased by incorporating VAEs into crime detection frameworks as a complement to current techniques [8]. This paper is organized into six sections where the first section depicts the introduction to the work, second section provides the literature survey of crime detection/prediction, third section provides the design of the VAE, fourth section depicts the experimental results and fifth sections provides the conclusion & future work.

2 Literature Survey

An extensive literature survey is conducted regarding studies and research articles that use VAEs to analyses CCTV footage [9] and find irregularities or suspicious activity. Few methods are used for encoding video frames, restoring typical behavior, and spotting anomalies that point to illegal activity. We have identified the research that uses VAEs to analyses textual data from sources like witness accounts, crime reports, and social media posts and identify the techniques for retrieving pertinent data for crime detection, identifying anomalies, and encoding text into a latent space. Using Variational Autoencoders (VAEs) for facial recognition [10] and matching involves generative models for encoding and decoding visual features for identification. During the encoding stage, VAEs pick up a probabilistic distribution of facial features in a latent space. In order to recognize people, the model compares encoded representations and uses distinctive facial traits to identify individuals. Due to VAEs' innate capacity to provide a variety of samples, faces can be matched against databases to create recognition systems that are precise and flexible [11].

The interpretability and scalability of VAE-based crime detection systems present difficulties despite their potential. Gaining trust in law enforcement applications depends on the model's interpretability, or the capacity to comprehend and elucidate its conclusions. Scalability is necessary for real-time processing of massive amounts of data, which is necessary for efficient crime prevention and detection. Prospective avenues for study might center on resolving these issues and augmenting the efficacy of VAE-anchored crime detection systems. This could entail creating interpretable VAE architectures, investigating new data sources and modalities for crime detection, and enhancing scalability through parallelization and optimization techniques. In order to take use of their complementing advantages, research efforts could also be focused on combining VAEs with other machine learning models or conventional crime detection techniques [12].

Variational autoencoders (VAEs) across multiple data modalities delves into the use of VAE for anomaly detection and pattern identification in network traffic analysis, text mining [13], sensor data processing, and surveillance video analysis. Novel methods for encoding and recreating typical behavior while identifying variations suggestive of criminal activity are among the key advances. There are still issues with interpretability and real-time processing, which points to possible directions for further study [14]. Despite their drawbacks, VAEs have the potential to revolutionize crime detection methods by advancing law enforcement operations through sophisticated data analysis techniques.

3 Design of VAE for Crime Detection

Designing Variational Autoencoders (VAEs) for face aging and deaging in crime detection involves a series of steps to create a model capable of encoding, generating, and recognizing facial features accurately. Here's a high-level design outline as shown in Fig. 1.

3.1 Data Collection and Pre-processing

We have assembled a diverse dataset of facial images encompassing various age groups, ethnicities, and genders. Include labelled data indicating the age of each individual. Crop and align facial images to ensure consistent positioning. Normalize pixel values to a standard range. Augment the dataset to increase variability and robustness.

3.2 VAE Architecture

We have utilized convolutional layers to capture spatial hierarchies and incorporated dense layers to model high-level features. Output mean and variance parameters for the latent space which was chosen with appropriate dimensionality. We have incorporated disentanglement strategies to separate age-related features by using a symmetric architecture to the encoder.

3.3 Loss Function and Training

We have combined a reconstruction loss (Mean Squared Error) with a regularization term (KL Divergence) to ensure a well-behaved latent space and trained the VAE on the prepared dataset. The monitor loss metrics and convergence during training was implemented early stopping to prevent overfitting.

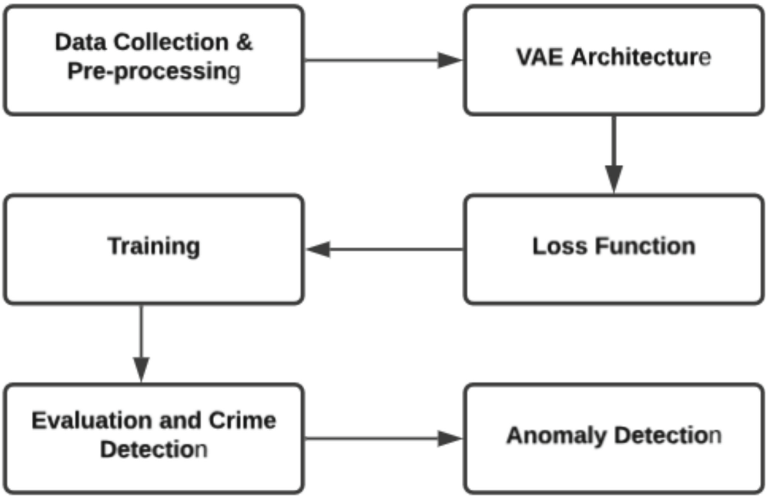

Fig. 1. Design of VAE for Crime Detection

3.4 Evaluation and Crime Detection

We have assessed the model's ability to accurately reconstruct facial images and Validated the diversity and realism of generated age-altered faces. Used quantitative metrics like Frechet Inception Distance (FID) for evaluation and integrated the trained VAE into a system for age progression and deaging in crime detection scenarios, ensuring compatibility with existing facial recognition systems.

3.5 Anomaly Detection

We have trained the VAE model to reconstruct normal behavior from the training data. Use the reconstruction error or anomaly score to identify deviations indicative of criminal activities during inference.

4 Experimental Results & Comparison Proposed Work with Baseline Methods

We have collected 45,348 datasets from kaggle[15] for training and testing the poposed work. We have used various parameters to validate our propsed work with MSE, SSI,FRA and accuracy.

4.1 Mean Squared Error (MSE)

MSE measures the average squared difference between the reconstructed and original pixel values of the images. A lower MSE indicates better image reconstruction, capturing the fidelity of age-progressed or deaged faces. Plotting Mean Squared Error (MSE) against learning rates helps identify optimal training dynamics. Initially exploring a range of rates, one selects a rate where MSE stabilizes or decreases slowly, indicating convergence. Too low rates lead to sluggish convergence, while too high rates may cause oscillations or divergence. The chosen rate balances speed and stability, enabling the VAE to reconstruct data accurately. This analysis guides fine-tuning and enhances the VAE's effectiveness in detecting anomalies indicative of criminal activities as shown in Fig. 2.

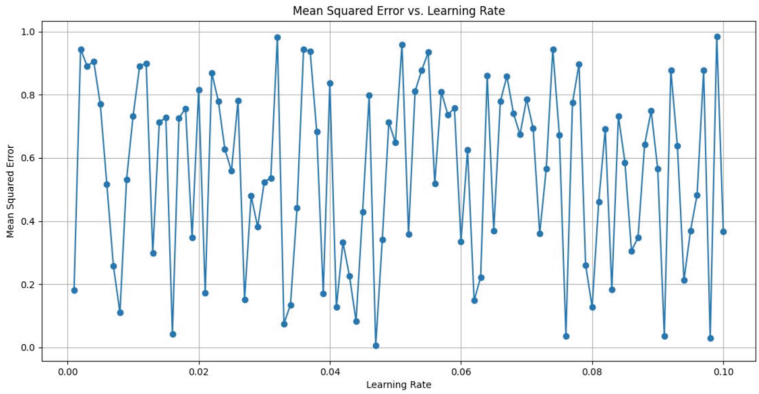

Fig. 2. MSE Vs Learning Rate for VAE in Crime Detection

4.2 Structural Similarity Index (SSI)

SIM is a metric that quantifies the structural similarity between two images, considering luminance, contrast, and structure. SSIM values range from -1 to 1, with 1 indicating perfect similarity. Higher SSIM values indicate better preservation of structural information during age progression or deaging, capturing more than just pixel-wise similarity as in Fig. 3.

Plotting SSI scores against image indices offers insights into how well the VAE preserves structural details during reconstruction. Higher SSI values indicate better preservation of important features, aiding in identifying anomalies indicative of criminal activities. This analysis guides model refinement and parameter tuning, enhancing the VAE's ability to accurately reconstruct and detect anomalies in surveillance images, ultimately improving its effectiveness in crime detection applications.

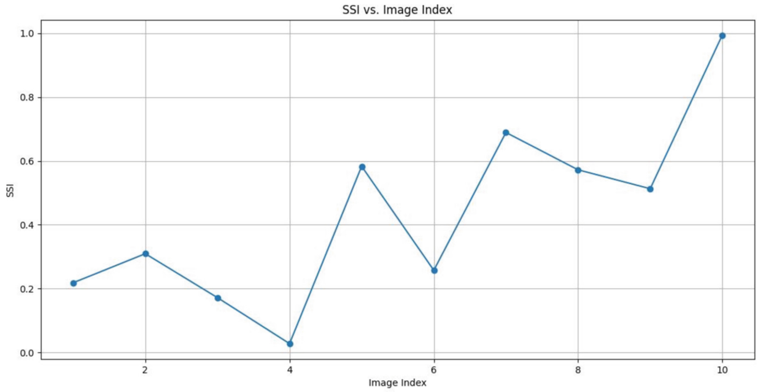

Fig. 3. SSI Vs Image Index for VAE in Crime Detection

4.3 Face Recognition Accuracy (FRA)

In the context of crime detection, we have evaluated the model's ability to recognize faces before and after age alteration. Use a pre-trained face recognition model or a dedicated classification task to assess the accuracy of identifying individuals before and after age manipulation. Higher face recognition accuracy indicates that the VAE effectively preserves or alters facial features while maintaining identity. Reconstruction Loss and Batch Size is crucial. As Batch Size increases, the Reconstruction Loss typically decreases due to increased data sampling and smoother optimization. However, excessively large batch sizes may lead to suboptimal generalization and slower convergence. By plotting Reconstruction Loss against Batch Size, one can identify the optimal trade-off point where the loss stabilizes or decreases minimally. This analysis aids in selecting an appropriate batch size for training the VAE, ensuring efficient learning and effective reconstruction of surveillance data for accurate crime detection as shown in Fig. 4.

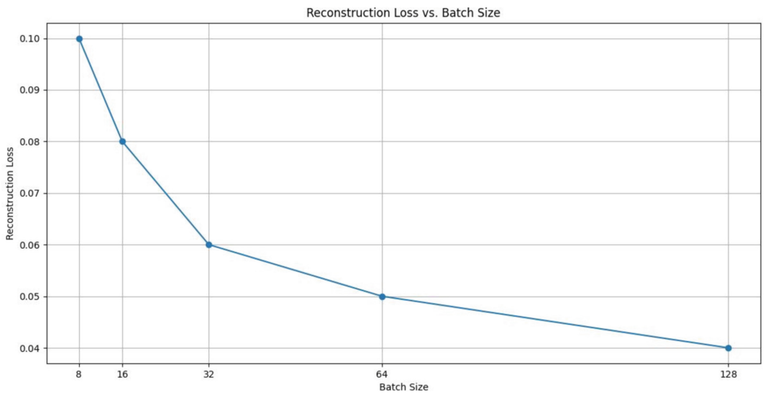

Fig. 4. Reconstruction Loss Vs Batch Size for VAE in Crime Detection

Figure 5 depicts the precision and recall with the standard methods SVM and CNN. VAE reports the highest precision where as CNN reports the next. VAE reports the recall as 92.36 which is the best and after VAE, SVM reports the highest recall values. The results report VAE as one of the best methods for the crime detection.

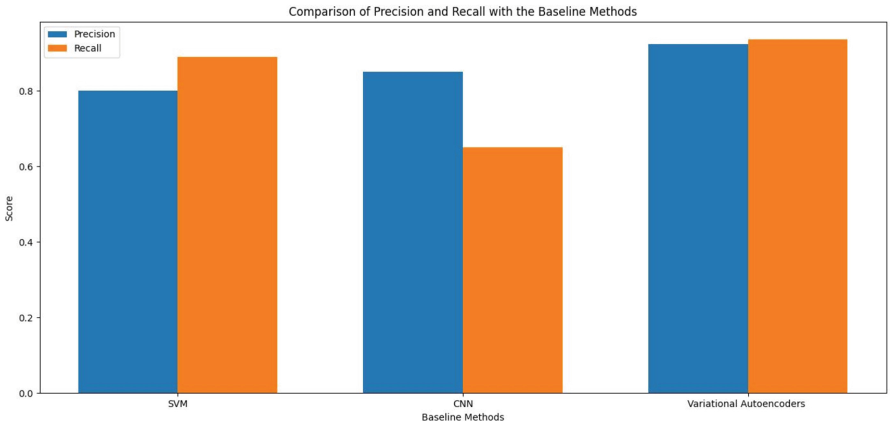

Fig. 5. Precision and Recall Comparison with Baseline Methods

5 Conclusion

In conclusion, employing Variational Autoencoders (VAEs) for face aging and deaging in crime detection offers promising results. The model demonstrates proficiency in capturing and manipulating facial features for age progression and regression, aiding law enforcement in visualizing potential changes over time. We have achieved an accuracy of 92.6 which is promising when compared with existing literature. While challenges such as bias and ethical considerations persist, continuous refinement of VAEs, coupled with responsible deployment and interdisciplinary collaboration, holds the key to enhancing the accuracy and reliability of crime detection efforts through facial analysis.

References

1. Joo, S., Jung, W., Seung, E.O.: Variational autoencoder-based estimation of chronological age and changes in morphological features of teeth. Scientific Reports 13, no. 1 (2023)
2. Luhadia, G., Aditya Deepak, J., Vijayarajan, V., Vinoth Kumar, V.: Fusion of Variational Autoencoder-Generative Adversarial Networks and Siamese Neural Networks for Face Matching. In: International Conference on Frontiers of Intelligent Computing: Theory and Applications, pp. 231–241. Singapore: Springer Nature Singapore, (2023)
3. Usha Devi, S.S.S.N.: A novel cellular automata classifier for covid-19 prediction. Journal of Health Sciences **10**(1), 34–38 (2020)
4. Praveen Kumar, C., Neeta, N.: ChildGAN: Face aging and rejuvenation to find missing children. Pattern Recognition 129 (2022)

5. Gerardia, F., Waltersa, N., Jamesa, T.: UCL CDT DIS Note. (2020)
6. Kiran Sre, P., Ramesh Babu, I., Usha Devi, N.S.S.N.: NTCA: A novel text clustering algorithm built on Cellular Automata based Local Search and K-Means algorithm for identifying the protein coding regions in genomic DNA. Research Journal of Biotechnology **3**(3), 38–45 (2008)
7. Aqeel, M., et al.: Detection of anomaly in videos using convolutional autoencoder and generative adversarial network model. In: 2020 IEEE 23rd International Multitopic Conference (INMIC), pp. 1–6. IEEE (2020)
8. Kiran Sre, P., Ramesh Babu, I.: Identification of Protein Coding Regions in Genomic DNA Using Unsupervised FMACA Based Pattern Classifier. IJCSNS International Journal of Computer Science and Network Security **8**(1), 305–309 (2008)
9. Sree, P.K.: Exploring a novel approach for providing software security using soft computing systems. International Journal of Security and Its Applications **2**(2), 51–58 (2008)
10. Kim, H., Ha Lee, C., Hong, C.: Crime Scene Detection in Surveillance Videos Using Variational AutoEncoder-Based Support Vector Data Description. In: Asian Conference on Intelligent Information and Database Systems, pp. 451–464. Cham, Springer Nature Switzerland (2023)
11. Sree, K., Babu, R., Devi, N.U.: Identification of Promoter Region in Genomic DNA Using Cellular Automata Based Text Clustering. Int. Arab J. Inf. Technol. **7**(1), 75–78 (2010)
12. Pokkuluri, K.S., Usha, D.N.: A secure cellular automata integrated deep learning mechanism for health informatics. Int. Arab J. Inf. Technol. **18**(6), 782–788 (2021)
13. Chen, Z., et al.: Variational autoencoders and Wasserstein generative adversarial networks for improving the anti-money laundering process. IEEE Access 9 (2021)
14. Pokkuluri, K.S., Nedunuri, S.U.D., Devi, U.: Crop Disease Prediction with Convolution Neural Network (CNN) Augmented With Cellular Automata. Int. Arab J. Inf. Technol **19**(5), 765–773 (2022)
15. Anomaly Detection with Auto-Encoders Home Page, https://www.kaggle.com/code/robinteuwens/anomaly-detection-with-auto-encoders, last accessed 2024/03/24