International Journal of Computational Science and Engineering Research
ISSN: XXXX- XXXX(Online) , http://www.ijcser.com/
Regular Issue , Vol. 2, Issue. 4 (October – December), 2025 , Pages: 1 - 7
Received: 15 May 2025 ; Accepted: 28 September 2025 ; Published: 06 October 2025
Original Paper : Applied Research – Comapre and Contrast Paper
https://doi.org/10.63328/IJCSER-V2I4P1

An Effective Cryptographic Algorithm For Multimodal Datasets Cryptanalysis Using Deep Learning

M. Pavan [1*] **, S. Dilli Babu** [2]

[1*] Master of Computer Application (MCA), , Mohan Babu University, Tirupati-517501; pavanm10949@gmail.com

[2] Associate Professor, Dept of CSE, SOC, Mohan Babu University, Tirupati ,India; dillibabusalvakkam@gmail.com

** Corresponding Author:* pavanm10949@gmail.com

Abstract: The cryptanalysis of multimodal datasets is particular, given the distinct statistics formats and cryptographic strategies that pose demanding situations to analysis. In this studies paintings, an green cryptographic algorithm for boosting the robustness and accuracy of cryptanalysis the usage of deep mastering is proposed. This method includes the usage of CNNs and transformer-based totally architectures in a systematic evaluation of multimodal datasets along with textual content, pictures, and numerical statistics to pick out the cryptographic styles and vulnerabilities. The model consists of feature extraction, modality fusion, and automated key inference to decode encrypted information with progressed efficiency.

Keywords: Kidney disease classification, CT images, CNN- LSTM, Deep learning, Grad-CAM, Medical imaging.

1. Introduction

Cryptanalysis is a cornerstone of modern cybersecurity, coping with the observe of cryptographic structures to apprehend their vulnerabilities and expand sturdy security measures [1,2]. With the upward push of deep learning and artificial intelligence (AI), conventional cryptanalysis has prolonged into multimodal datasets, combining numerous types of facts consisting of text, pics, and numerical streams [3,4]. These datasets are of amazing importance for growing state-of-the-art cryptographic algorithms and comparing their resilience against emerging threats [5,6]. However, building and analyzing those datasets gift a few challenges concerning records exceptional, annotation, and processing that have an immediate have an effect on on device studying (ML) version overall performance [7]. Machine Intelligence, especially inside the form of deep studying, guarantees to accelerate cryptanalysis and to become aware of and examine cryptographic algorithms at extraordinary scales and accuracies [8,9].Such ML fashions require massive representative datasets for education, validation, and checking out. Advanced annotation techniques may be required for the complexity of cryptographic datasets on the way to capture subtle relationships and patterns [10]. Furthermore, these datasets should be curated to decrease errors and biases that might propagate thru ML fashions, undermining their reliability and accuracy [11,12]. They require the synthesis of more than one statistics stream to pick out correlations and styles throughout modalities [13,14]. For cryptanalysis, this can involve linking encrypted textual content, metadata, and associated visual or numerical information to uncover vulnerabilities or validate algorithmic robustness [15]. Annotation equipment are important in this process, permitting researchers to label and shape facts correctly, enabling ML fashions to analyze from rich, multimodal sources [16]. Since the creation of superior deep studying fashions, which includes Transformers, by way of Vaswani et al. In 2017 [17], foundational improvements in natural language processing, computer vision, and multimodal evaluation have become a truth [18]. Transformer-based totally architectures, with their in-context learning and scalability, provide a effective framework for cryptanalysis duties, permitting structures to adapt and generalize with minimum human intervention [19].

2. Literature Review

O.Al-Hamdani et al. [17] designed a reliable multimodal system for biometric verification incorporating speech, Electrocardiogram (ECG), and Phonocardiogram (PCG) signals. It was shown that features extracted from heart and speech signals using Mel Frequency Cepstral Coefficients (MFCC) have superior performance in multimodal fusion with simple sum score fusion and piecewise-linear normalization. However, they highlighted that the variation of the

ISSN: XXXX- XXXX(Online) , http://www.ijcser.com/ , Vol. 2, Issue 4, 2025, https://doi.org/10.63328/IJCSER-V2I4P1

heartbeats depending on emotional conditions such as being joyful or being stressed could be reflected in the ECG signal, thereby making biometric analysis more complex.

M.S.M. Asaari et al. [18] presented the fusion of finger vein and finger geometry recognition for biometric identification. By using Band-Limited Phase-Only Correlation (BLPOC), they obtained enhanced accuracy in the recognition of finger geometry compared to single modality recognition. However, the non-linear distortions present in finger-vein images resulted in a degradation of performance in BLPOC-based matching, emphasizing the need for the consideration of image quality in multimodal biometric systems.

M.S. Aslan et al. [19] applied Auto Encoders (AE) in multi-channel, multi-modal feature learning to face recognition. In the presented work, an AE, in collaboration with ADMM, is applied to optimize the extraction of the features as well as reduce the energy consumption through a task decomposition onto sub-bands. Although their method was promising toward the integration of multi-modalities, the authors underscored that applying CNNs to these models would require a very large number of samples to avoid overfitting, which was difficult in real applications.

N. Radha and A. Kavitha [20] proposed a multimodal biometric system that employed iris and fingerprint recognition. In this system, rank-level fusion was used for the combination of biometric features and Fisher Linear Discriminant (FLD) for feature extraction. This led to an improved accuracy of the system, but the authors of the paper realized that there existed an inherent rank aggregation problem. This could be problematic in fusion when the feature rankings were not matched.

A. Jagadeesan et al. [21] proposed a multimodal biometric system for constructing an iris and fingerprint-based secure cryptographic key. The methodology represented the feature set formed from minutiae point extraction in the iris and texture features in the fingerprint, thus having a 256-bit length that was secure as a cryptographic key for use in biometric-based authentication systems, BIS. The performance metrics were found to be good with both security and accuracy. However, it depends on low-level cryptographic techniques, which will not withstand advanced attacks.

Based on deep neural networks, it learns optimal features from multimodal datasets, yielding higher recognition rates and improved mechanisms of fusion. Furthermore, for security vulnerabilities seen in previous systems, the system uses advanced cryptography methods such as RSA (Rivest-Shamir-Adleman) for key generation. In order to overcome some of the present limitations of the traditional multimodal biometrics, deep learning-cryptographic algorithm fusion can bring about a bright perspective for more secure and efficient application in areas of access control, verification of identity, etc.

Another reason why deep learning methods, such as Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs), have gained significant attention is due to their capacity to learn hierarchical representations of multimodal data with both improved accuracy and robustness. Deep learning methods are the best at extracting features from high-dimensional complex data, which is very suitable for multimodal biometrics that comprise diverse data types, such as images, speech, and ECG signals. These models along with cryptographic key generation systems potentially result in highly secure and scalable biometric authentication solutions.

3. Methodology

The methodology for cryptanalysis of cryptographic algorithms in multimodal datasets using deep learning starts with the collection and preprocessing of multimodal data, including textual, visual, and auditory information. First, the datasets are cleaned, normalized, and transformed into suitable forms for deep learning models. Feature extraction techniques, including NLP for textual data, CNNs for image data, and RNNs for audio data, are applied to derive high-level representations of the different modalities. The combined features are then structured into a unified vector form that represents the multimodal data comprehensively. This dataset is used to train a deep neural network model capable of learning complex patterns and relationships between the data types.

Once the multimodal features are extracted, a hybrid deep learning architecture is used for the cryptanalysis process. The architecture may be a combination of CNNs, RNNs, and fully connected layers to leverage the spatial, sequential, and contextual information from different modalities. This is trained on cryptographic cipher texts and their corresponding plaintexts that are mapped to the way encryption schemes would modify the original data. The model performance is tested using the proposed algorithms. This includes AES or RSA, for example, during training to check for vulnerabilities.

HOW WORKS CNN & RNN

Convolutional Neural Networks (CNNs): CNNs are designed specifically for image processing type of grid data. They employ three main types of layers: convolutional, pooling, and fully connected layers. Convolutional layer applies filters to input images to generate feature maps which

ISSN: XXXX- XXXX(Online) , http://www.ijcser.com/ , Vol. 2, Issue 4, 2025, https://doi.org/10.63328/IJCSER-V2I4P1

highlight important features such as edges and textures. Mathematically, this involves convolution operations where a filter matrix

Recurrent Neural Networks (RNNs): RNNs are used for sequential data, such as time series or text. In contrast to feedforward neural networks, RNNs have a hidden state that captures information from previous time steps, which enables them to model dependencies over time

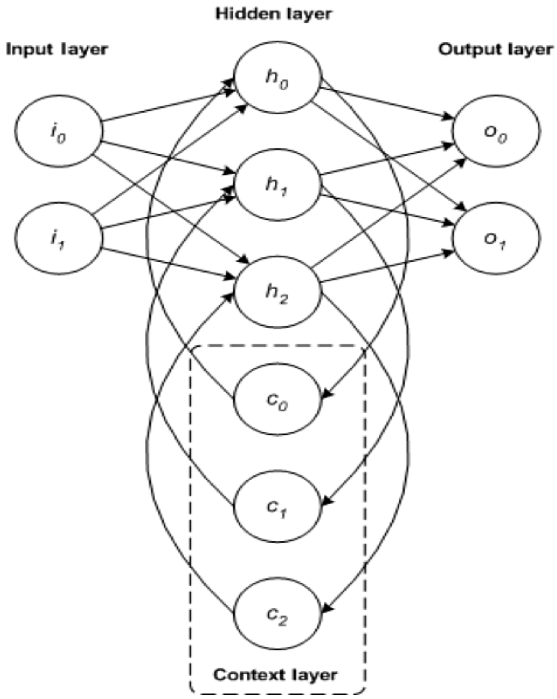

Figure.1 Recurrent Neural Networks (RNNs)

4. Problem Definition

In the changing virtual international, stable conversation relies upon on cryptographic algorithms to shield touchy statistics. However, the increasing use of multimodal datasets, which encompass various statistics sorts including textual content, picture, and audio, creates a brand new challenge in cryptanalysis—the take a look at of reading and breaking cryptographic systems.

Traditional methods of cryptanalysis are not able to successfully pick out vulnerabilities or classify cryptographic algorithms in such complex datasets due to the range and complexity of facts structures.

This trouble is even more extreme in programs which include cybersecurity, digital forensics, and management of encrypted facts, wherein multimodal datasets are not unusual. Additionally, guide analysis is time-ingesting, errors-inclined, and regularly fails to discover hidden styles or correlations crucial to information cryptographic behaviours.

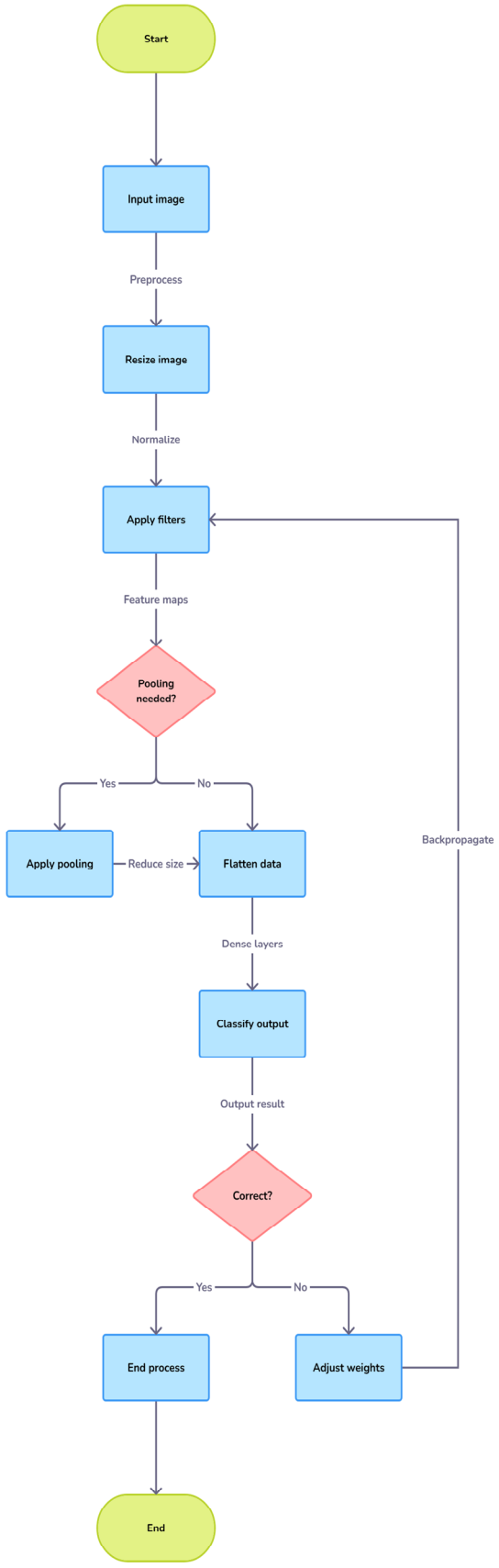

Figure.2 Convolutional Neural Networks (CNN)

5. Existing System

At the instant, maximum detection and cryptanalysis strategies used are based totally on

ISSN: XXXX- XXXX(Online) , http://www.ijcser.com/ , Vol. 2, Issue 4, 2025, https://doi.org/10.63328/IJCSER-V2I4P1

traditional strategies that include statistical evaluation, pattern recognition, and rule-based totally structures. These structures tend to be a venture whilst used with multimodal datasets because they have got complicated, high-dimensional records. Here's an outline of the existing systems used for cryptographic set of rules detection and cryptanalysis in multimodal contexts:

5.1. Machine Learning-Based Cryptanalysis

Machine gaining knowledge of techniques together with supervised and unsupervised gaining knowledge of have additionally been used extra recently in cryptanalysis. Some common techniques include:

Support Vector Machines (SVMs): They classify whether or not a given information pattern is encrypted or not, depending at the sample observed inside the ciphertext.

Decision Trees: Used to classify and are expecting the kind of cryptographic structures from the attributes of the ciphertext and other observable features.

Deep Neural Networks (DNNs): More complex type and regression responsibilities. These fashions can seize high-dimensional relationships inside the information, but they require large datasets and extensive computational sources to teach efficaciously.

5.2. Cryptanalysis of Multimodal Datasets

Analyzing multimodal statistics (e.G., combining image, textual content, and audio statistics) in cryptanalysis is extra complex due to the form of information codecs and systems. Existing procedures for multimodal cryptanalysis encompass

5.3. Deep Learning in Cryptographic Algorithm Detection

Deep learning techniques have great potential in enriching the detection and analysis of cryptographic algorithms, especially complex multimodal data:

Convolutional Neural Networks (CNN): Mainly adopted for image-oriented cryptanalysis; it is found useful for unveiling visual patterns contained in the images that are already encrypted or it may identify specific watermarking as well as steganography-based schemes.

Recurrent Neural Networks (RNNs): This category, again applied over sequential data text or speech contains LSTMs and it proved useful in such areas of determining the cryptographic actions, padding/ or key-schedules contained into the encrypted streams of text / audio.

6. Proposed System

6.1. Multimodal Dataset Collection and Preprocessing:

Data Types: The system is designed to support multiple types of data (text, images, audio, etc.). Datasets may originate from various sources such as encrypted files, network traffic, or multimedia files.

Data Preprocessing: Every modality of data, be it text, image, or audio, is pre-processed in order to be converted into an appropriate format for deep learning models:

Text: Tokenization, padding, and vectorization, for example word embeddings or TF-IDF.

Images: Scaling, whitening, and padding (if necessary).

Audio: Feature extraction (e.g., spectrograms, MFCCs).

6.2. *Cryptographic Feature Extraction:*

Transformation Analysis: The system will subject the encrypted data for possible cryptographic transformations like substitution and permutation, key management, and patterns from encryption algorithms.

Text: Recurrent Neural Networks (RNNs), Long Short-Term Memory (LSTM), or Transformer models to capture sequential dependencies.

Images: Convolutional Neural Networks (CNNs) will be used to detect patterns and transformations in encrypted images.

Audio: A combination of CNN and RNN to detect cryptographic transformations in audio signals.

6.3. Deep Learning Model Architecture

Multi-Modal Neural Network: A multi-input neural network architecture capable of processing different modalities in parallel will be used. The architecture could include:

Model Training: The model will be trained using a supervised or semi-supervised approach based on the availability of labeled data. If labeled data is not available, then unsupervised techniques such as autoencoders or Generative Adversarial Networks (GANs) may be used for anomaly detection.

6.4. Cryptanalysis

Classification Layer: After feature extraction and fusion, a classification layer will identify the cryptographic

ISSN: XXXX- XXXX(Online) , http://www.ijcser.com/ , Vol. 2, Issue 4, 2025, https://doi.org/10.63328/IJCSER-V2I4P1

algorithm used. The system will be trained to classify a variety of cryptographic algorithms, including symmetric and asymmetric encryption methods (AES, RSA), and hashing algorithms (SHA-256, MD5).

Model Optimization: Hyperparameter tuning, Transfer Learning, and Fine-tuning on pre-trained models are going to be applied for enhancing the accuracy of the system.

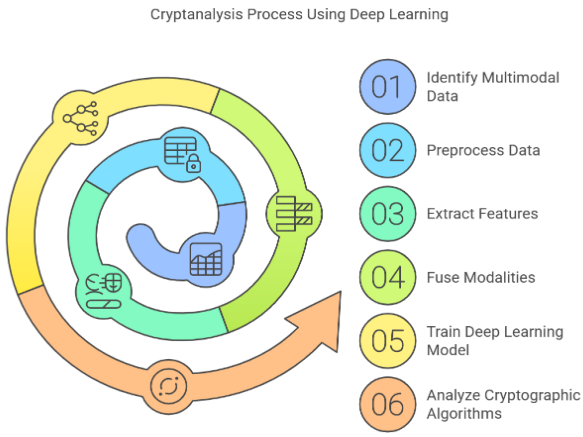

Figure.3 Cryptanalysis process using Deep Learning

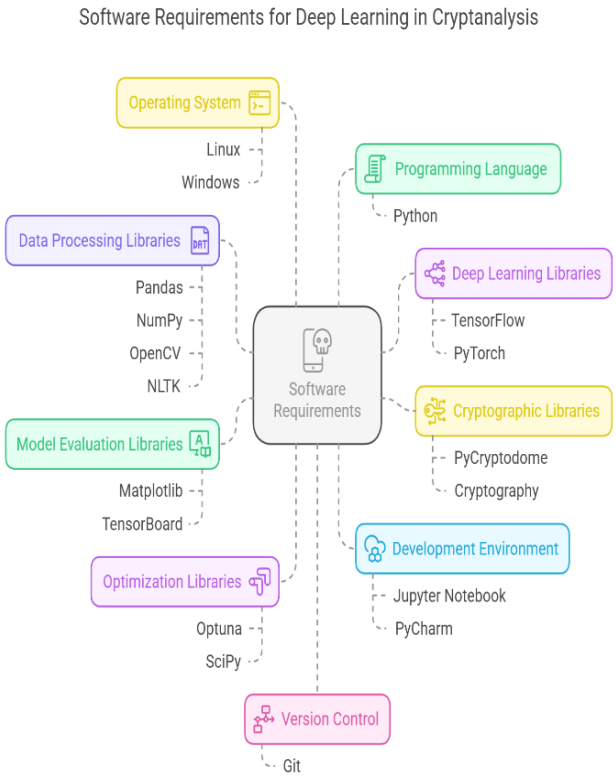

Figure.4 Software Requirements for Deep Learning in Cryptanalysis

7. Problem Formulation

Given a multimodal dataset D={D1,D2,…,Dn}D = \{D_1, D_2, \dots, D_n\}D={D1,D2,…,Dn}, where each DiD_iDi represents a distinct type of data (such as text, image, or audio), we are tasked with solving the following:

7.1. Cryptographic Algorithm Design:

Let AAA represent a cryptographic algorithm applied to the dataset. The encryption of data DiD_iDi with AAA results in a transformed encrypted dataset Ei=A(Di)E_i = A(D_i)Ei=A(Di). The cryptographic algorithm AAA should meet the criteria of being resistant to cryptanalysis on the specific data types.

7.2. Deep Learning Model for Cryptanalysis:

The cryptanalysis model Mcryptanalysis M_{\text{cryptanalysis}}Mcryptanalysis uses a deep learning framework (such as a convolutional neural network or recurrent neural network) to identify the encryption patterns in the multimodal data.

The objective is to develop a model McryptanalysisM_{\text{cryptanalysis}}Mcryptanalysis that can take an encrypted dataset EiE_iEi and predict the underlying structure or weaknesses in AAA, represented as Mcryptanalysis(Ei)≈DiM_{\text{cryptanalysis}}(E_i) \approx D_iMcryptanalysis(Ei)≈Di.

7.3. Multimodal Fusion

Fusion techniques are needed to combine information from different data modalities. This can be achieved by employing hybrid models that learn joint representations for multimodal data. Let F\mathcal{F}F represent the fusion function that combines different modality representations. The combined representation can be expressed as: F(D1,D2,…, Dn)=F(D1,D2,…,Dn)F(D_1, D_2, \dots, D_n) = \mathcal{F}(D_1, D_2, \dots, D_n)F(D1,D2,…,Dn)=F(D1,D2,…,Dn)

7.4. Methodology Of The Proposed Algorithm

7.4.1. Data Gathering and Preprocessing

Multimodal Data Collection: Collect a multimodal dataset that contains the encrypted data along with its ciphertext. The dataset may contain any type of data, such as text, images, or time-series data, to test the algorithm's performance in different domains.

Data Preprocessing: Clean and preprocess the statistics into right format. This ought to include normalization, characteristic extraction, and encoding. For text information, techniques together with tokenization and vectorization, using phrase embeddings like Word2Vec, GloVe might be in use. For photograph data, techniques

ISSN: XXXX- XXXX(Online) , http://www.ijcser.com/ , Vol. 2, Issue 4, 2025, https://doi.org/10.63328/IJCSER-V2I4P1

like resizing and making use of histograms may be in exercise.

7.4.2. Model Architecture

Deep Learning Model: Design a deep neural network (DNN) architecture that could procedure multimodal inputs. For instance, a combination of CNNs for image facts and Recurrent Neural Networks (RNNs) or Transformer models for textual data.

CNN for Image Cryptanalysis: The use of CNN is to become aware of the hidden structures or visual patterns inside encrypted images. In the context of cryptanalysis, CNN could be trained to find encryption artifacts or cipher styles.

7.4.3. RNN or Transformer for Text Cryptanalysis

For the dataset of encrypted texts, the technique can contain using RNN or Transformer-based fashions inclusive of BERT for linguistic sample reputation and relationships a number of the ciphertext factors.

7.4.4. Training the Deep Learning Model

Dataset Splitting: Split the dataset into schooling, validation, and testing units to assess version performance nicely.

Model Training: The deep gaining knowledge of model is trained via the use of techniques like backpropagation and gradient descent. A appropriate loss function for measuring the accuracy of the prediction, such as go-entropy in type tasks and Mean Squared Error (MSE) in regression, is implemented.

7.4.5. Cryptanalysis and Evaluation

Cryptanalysis Process: Once the deep learning model is trained, it should be able to recognize cryptographic algorithms or weaknesses in the encrypted multimodal datasets. The model can predict: The type of cryptographic algorithm used.Whether the data is encrypted or not.A decryption key or a partial decryption of the data if applicable.Performance Evaluation: This should be assessed on a different scale of performance metrics, like accuracy, precision, recall, F1 score, and even Area Under Curve (AUC) for the classification task and for regression task, it will be RMSE.

8. Conclusion

This study recognizes the increasing complexity of obtaining multimodal datasets, which are usually comprised of heterogeneous data types like text, images, audio, and video. The research seeks to create a secure cryptographic algorithm specifically designed for the specific challenges presented by such datasets, providing improved security and privacy. In addition, it utilizes deep learning methods to conduct cryptanalysis, allowing for the detection of possible vulnerabilities and the fortification of cryptographic systems. This project recognizes the work of researchers in both deep learning and cryptography whose ground-breaking work has led to breakthrough solutions in data security. By merging these fields, the project aims to tackle the changing threats in the digital world, securing sensitive data in a world of increasingly complex cyberattacks.

References

[1] A. Karchmer, "On Stronger Computational Separations Between Multimodal and Unimodal Machine Learning," arXiv preprint arXiv:2404.02254, 2024.

[2] V. Talreja, M. Valenti, and N. Nasrabadi, "Deep Hashing for Secure Multimodal Biometrics," arXiv preprint arXiv:2012.14758, 2020.

[3] V. Talreja, S. Soleymani, M.C. Valenti, and N.M. Nasrabadi, "Learning to Authenticate with Deep Multibiometric Hashing and Neural Network Decoding," arXiv preprint arXiv:1902.04149, 2019.

[4] M.S. Aslan, Z. Hailat, T.K. Alafif, and X.W. Chen, "Multi-channel multi-modal feature learning for face recognition," Pattern Recognition Letters, vol. 85, pp. 79-83, 2017.

[5] G.W. Mwaura, W. Mwangi, and C. Otieno, "Multimodal Biometric System: Fusion of Face and Fingerprint Biometrics at Match Score Fusion Level," International Journal of Scientific & Technology Research, vol. 6, no. 4, pp. 41-49, 2017.

[6] Y. Bengio, A. Courville, and P. Vincent, "Representation learning: A review and new perspectives," IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol.35, No.8, pp.1798-1828, 2013.

[7] I. Goodfellow, Y. Bengio, and A. Courville, "Deep Learning," MIT Press, Vol.2016, 2016.

[8] H. H. Yang and H. W. Chan, "Cryptanalysis of machine learning models: An emerging research area," ACM Computing Surveys, Vol.52, No.1, pp.1-30, 2020.

[9] A. K. Jain, P. Flynn, and A. A. Ross, "Handbook of Biometrics," Springer Science & Business Media, Vol.2007, pp.1-49, 2007.

[10] R. Shokri and V. Shmatikov, "Privacy-preserving deep learning," In: Proceedings of the 22nd ACM SIGSAC Conference on Computer and Communications Security, pp.1310-1321, 2015.

[11] T. Salimans et al., "Improved techniques for training GANs," Advances in Neural Information Processing Systems, Vol.29, pp.2234-2242, 2016.

[12] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," Advances in Neural Information Processing Systems, Vol.25, pp.1097-1105, 2012.

ISSN: XXXX- XXXX(Online) , http://www.ijcser.com/ , Vol. 2, Issue 4, 2025, https://doi.org/10.63328/IJCSER-V2I4P1

[13] M. Abadi et al., "TensorFlow: A system for large-scale machine learning," In: Proceedings of the 12th USENIX Symposium on Operating Systems Design and Implementation, pp.265-283, 2016.

[14] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp.770-778, 2016.

[15] S. Hochreiter and J. Schmidhuber, "Long short-term memory," Neural Computation, Vol.9, No.8, pp.1735-1780, 1997.

[16] J. Brownlee, "Machine Learning Mastery with Python," Machine Learning Mastery, Vol.2016, 2016.

[17] S. Ruder, "An overview of gradient descent optimization algorithms," arXiv preprint arXiv:1609.04747, 2016.

[18] R. Vaishya, M. Javaid, I. H. Khan, and A. Haleem, "Artificial Intelligence (AI) applications for COVID-19 pandemic," Diabetes & Metabolic Syndrome: Clinical Research & Reviews, Vol.14, No.4, pp.337-339, 2020.

[19] X. Wu et al., "Data mining with big data," IEEE Transactions on Knowledge and Data Engineering, Vol.26, No.1, pp.97-107, 2014.