Hiding Financial data in applications based on AES Encryption and Steganography Method

MSc Research Project
MSC in Cyber Security Evening-Year 2- MSCCYBE_JANOL23

Parwat Lohani
Student ID: x22179232

School of Computing
National College of Ireland

Supervisor: Michael Pantridge

National College of Ireland

MSc Project Submission Sheet

School of Computing

Student Name: Parwat Lohani ……………………………………………………………………………………………………

Student ID: x22179232……………………………………………………………………………………

Programme: MSC in Cyber Security **Year:** 2023-2024

Module: Academic Internship………………………………………………………………………

Supervisor: Michael Pantridge………………………………………………………………………

Submission Due Date: 12th Aug 2024…………………………………………………………………….…

Project Title: Hiding Financial data in applications based on AES Encryption and Steganography Method………………………….

Word Count: 5678…………………….…. **Page Count:** 18…………………………..…….

I hereby certify that the information contained in this (my submission) is information pertaining to research I conducted for this project. All information other than my own contribution will be fully referenced and listed in the relevant bibliography section at the rear of the project.
ALL internet material must be referenced in the bibliography section. Students are required to use the Referencing Standard specified in the report template. To use other author's written or electronic work is illegal (plagiarism) and may result in disciplinary action.

Signature: Parwat Lohani……………………………………………………

Date: 12th Aug 2024……………………………………………………

PLEASE READ THE FOLLOWING INSTRUCTIONS AND CHECKLIST

Attach a completed copy of this sheet to each project (including multiple copies)	□
Attach a Moodle submission receipt of the online project submission, to each project (including multiple copies).	□
You must ensure that you retain a HARD COPY of the project, both for your own reference and in case a project is lost or mislaid. It is not sufficient to keep a copy on computer.	□

Assignments that are submitted to the Programme Coordinator Office must be placed into the assignment box located outside the office.

Office Use Only	
Signature:	
Date:	
Penalty Applied (if applicable):	

Hiding Financial data in applications based on AES Encryption and Steganography Method

Parwat Lohani
x22179232
MSC in Cyber Security Evening-Year 2- MSCCYBE_JANOL23
National College of Ireland

Abstract

As per NASDAQ Financial crime report, $465.8B are lost to financial crimes in 2023 and organizations are trying to secure and protect their customers data like password, credit card more than ever. With the advancement in hardware and computing infrastructure it is easier to crack encrypted data, hence, this research Paper proposed a detailed implementation of cryptography and steganography together to store customer data without compromising Confidentiality, Integrity and image quality and making it difficult to extract and retrieve. A model was created using Python and tested through METLAB on the public images data set provided by University of Southern California. The approach was, *first*, encrypt, using AES-256 Bit, the various payloads, ranging from 10K bytes to 80K bytes and ,*second*, embedding the Payload in the images. The efficacy of the model was tested by analysing the PSNR, MSE, NCC, BER values. The tests showed a decrease in PSNR, indicating successful data hiding and An increase in NCC confirmed that the steganographic images remained visually similar. Based on the results captured from the tests, it can be concluded that approach mentioned in this research paper holds significant potential in the financial market for concealing customer data.

Keywords: *Cryptography, Cipher Block Chaining, AES, Steganography, Peak-Signal-to-noise ratio (PSNR), Mean Squared Error (MSE), Normalized Cross-Correlation (NCC), Bit-Error-Rate(BER)*

1. Introduction

While making online payment or creating a profile we provide our Personnel identifiable information (PII) e.g. name, date of birth, credit card etc., which travels through secure network (using TLS 1.3 etc.). These details are susceptible to hacking and as per 2024 CISO report [18] the spectrum and volume of attacks has been on rise across all categories.

To safeguard the financial data organizations implementing security at 2 levels, as shown in Figure 1, first, encrypt the data at-rest using Advanced Encryption standard like AES-CBC, 256 bit encryption. Second, encrypt the data in-transit using secure socket layer, 3D secure(by Visa); these techniques are discussed in details by Wei-Hsun, et al.,2018[16] , in his research paper.

In case of level first, While AES- 256 bit encryption is the strongest and robust encryption available today, however, it has some drawbacks as well as defined below.

1. Vulnerable to Paddling oracle attack, which allow decrypt ciphertext block by exploiting vulnerabilities in padding scheme. e.g. CVE-2019-1559: This vulnerability allows attackers to recover the plaintext of encrypted messages by exploiting a padding oracle in the implementation of AES-CBC.
2. CBC mode susceptible to various cyberattack
    a. Timing attack/Lucky13 attack: where attack measure time it takes for the server to decrypt to makes sure padding is valid or not, allowing attacker to decrypt ciphertext
3. CBC provide confidentiality but doesn't provide integrity protection. though the recommendation is to use authenticated encryption modes like AES-GCM or AES-CCM instead of AES-CBC. However, for organization using AES-CBC, appropriate countermeasure is needed.

In case of level second, though the data in-transit is encrypted however, if network is breached, this will lead to PII of customers & users going into wrong hands, which not only creates compliance issues, *e.g. GDPR*, but also damage organization's reputation. As per Balasubramanian, and S.Geetha, 2014 [13, To meet the Security objectives of Confidentiality, Integrity, Availability and Accountability organizations are using various methods MFA/3DSecure/TLS, However, they are also having drawbacks as they are vulnerable to attacks, expected benefit outnumbered the gains, high cost etc.

As per NASDAQ Financial crime report[21], $465.8B are lost to financial crimes in 2023 and organizations are trying to secure and protect their customers data like password, credit card more than ever. With the advancement in hardware and computing infrastructure it will be easier to crack encrypted data. Organizations can use steganography to hide confidential data in images however, using only steganography poses many vulnerabilities like steganalysis attack, frequency domain attack, deep learning based attacks e.g. CVE-2023-48022 – exploited open-source framework Ray to access source code, credentials and cloud access tokens.

Contrary to the traditional way of steganography using LSB, to safeguard the customer's critical financial data, this paper recommends a consolidated approach to safeguard financial data using cryptography AES-CBC encryption and Steganography using Transform domain method by hiding information within images without compromising confidentiality, integrity and image quality and making hackers difficult to extract and retrieve the financial data hence meeting security objectives. This method can also be implemented on files which contains financial data e.g. daily settlement and clearing report which Banks send to regulators.

The structure of research paper starts with

- Literature review, which consists of review of important research papers in the fields of Cryptography and steganography, outlining the objective, contrasting arguments and conclusion, followed by reasoning for why proposed method has been choses.
- The next sections describe the design consideration of proposed solutions and its implantations including the framework, tools and steps performed.
- The last section describes evaluation, which includes, comprehensive analysis of the results and main findings of the study and how the current implementation can be taken forward for future work.

2. Related Work

Various research and studies have been done, and different approaches also have been developed in the areas related to safeguarding financial data. The following section focussed on determination of the various aspects that are associated with the fields of Cryptography and Steganography their approach and various implementation techniques that have already been conducted.

2.1. Usage of Steganography to avoid SQL injection by saving credential hash in image.

In George, Thomas Ronil, 2018[6], research paper discussed about OWASP 2018 Top 1 vulnerability, SQL Injection, which exploits applications based on user provided inputs. The objective of authors research is to identify methods to safeguard the user credentials, e.g. user id and password, by storing data in images using steganography. However, with the modern tools like john the ripper or rainbow table hackers can not only identify the encryption used Md5, SHA126/256/512 but also extract the credentials.
The author discussed multiple implementations of password cracking techniques, first, *as per Wu et al. (2011)* [2], by optimizing the reversal of Md5 hash using GPU's parallel architecture CUDA (Compute unified device architecture). Second, *by Murakami et al. (2010)* [10], using John the ripper password cracking tool with dictionary file; these methods created as many threads number of passwords in the dictionary file hence each thread focus only on 1 hash to reverse. Hence the author postulate that with enhancement of hardware infrastructure and processing power no password is Safe.

The author deduced the research, conduction by Morkel et al. (2005) [17] & Wang and Wang (2004)[12], which provides different methods of hiding information within information using Image steganography. The researchers recommended to use least significant bit (LSB) of Spatial domain method of Steganography to store data in the last bit of each color. The last or redundant bits do not make any difference even if altered hence data can be kept hidden in these bits. The author proved the use of LSB method to hide the hashed credentials in images by implementing an application written in Python. The Implementation consists of creation of 2 separate tables and Python scripts. The research successfully showed password hash stored in image securely and retrieved. This technique provides additional security to password hashed. The concept of storing data securely can be extended to various industry segments like Banking and financial sectors and approach suggested of using LSB is key contributor to my research solution.

However, even if the credentials are hashed and stored in database, they are vulnerable to hacking; two reasons, *first*, hashes are easier to crack, *second*, efficiency of steganography depends on non-modification of solid color, saturated portions and if the payload is more then it is difficult to maintain image quality.

To provide extensive protection to data, the next section talks about data hiding using both cryptography and steganography.

2.2. Proposed system for data hiding using cryptography and steganography.

Sarmah, Dipti Kapoor et.al., 2010 [7], research paper introduced more robust data hiding secure technique using cryptography AES and Steganography using spatial method together. where in it is difficult to detect message in Frequency domain using DCT, Wavelets and FFT transformation. These methods are widely used together or independently by industries; However, it is possible for intruder to detect and decipher the message which leads to confidentiality and integrity issues.

The author proposed two-step approach, *first,* to use AES algorithm-*due to advantages like Very secure, cost effective, simple and flexible* - where in as per Rijndael algorithm [11] , a symmetric cipher can process 128bit block using 128 or 192 or 256 bits cipher keys. The algorithm process input by encrypting with keys and give ciphertext as output with a constrain that both input and output should be of same length and in AES only 128-bit length is allowed. The author recommends to Cipher blocker chaining (CBC) as it is considered more secure due to initialization vector and XORed.

Second, for steganography, Frequency domain algorithm (DCT) provides various advantages like secure, flexible, hard to detect and availability of different techniques for manipulation of Discrete Cosine transform DCT coefficient values explained by Huaiqing wang and Shouzhong wang in their 2004 [12] research paper. The author combined AES encryption and DCT techniques and implemented the solution to hide and retrieve data using VC++ 6.0 platform and free license libraries Arisimage Routines and Cximage599c. Based on the described method, *first*, the message is encrypted using AES using 128bit key and cipher text is created. *Second*, two extra keys are generated, *third*, hide the cipher text in the image by scrambling the alphabets in ciphertext and alter the DCT and inverse it to make Stego image. Inverse was done to retrieve the data from the image.

Although usage of cryptography and Steganography is considerably secure in hiding and retrieving of data using two more keys makes the application robust however, key management is an issue. Because for financial (credit card, PCI-data etc.) and compliance (GDPR data) the approach of creating ever changing keys real-time and storing it in secure place to decipher may pose vulnerability.

Therefore, in the next section the author is using various methods of cryptography and steganography and avoiding key management issues.

2.3. Robust steganography: Hiding financial data using contrast level value and text encryption.

Shafry et. al, 2020[8], research paper proposed using LSB and DCT along with encrypting and storing customer secret data (CSD) with segmentation encryption text method (SETM) in images and identify the efficacy and quality of image using Peak-signal to noise-ratio (PSNR)% which is a key requirement.

The author discussed about the financial data, especially e-transaction or online data which transmit across networks through multi-channel mode like B2B, B2C, C2C, client to bank etc., and associated cryptography and Steganography for securing and protecting the data. As per a separate research, Madhu, et. al., 2019[14], mentioned that Steganography stores the data either in reversible or irreversible mode based upon if overwritten LSB is recoverable or not; and recommends that irreversible mode offers more protection by providing greater embedding capability.

Also, Divya.A, S.Thenmozhi, 2016[1] did a detailed comparative analysis of various Steganography techniques in Spatial and Transform domain. Each of the methods in these domains has their own advantages and dis-advantages and limitation. Spatial domain uses techniques called LSB substitution to directly modify last bits from 24 bits pixel.

Also, Halder R, Sengupta , 2016 [5] proposed usage of HASH function on LSB and RSA algorithm to encrypt the data. In this approach position of LSB to hide secret data is determined by hash function, once LSB is know the secret message is embedded and hash is returned.

Though these methods rely on direct substitution at LSB and they are simple and effective to store financial data however chances of detection and extraction are higher. Moreover, with the technological advancement in the area of hardware and CPU optimization, as mentioned by Wu, H., Liu, X. and Tang, W., 2011[2] in their research paper, where in reduction of MD5 reverse steps from 64 to 49 increased the speed up to 16 times.

On the other hand, transform domain techniques like Discrete Cosine Transform (DCT) , Curvelet Transform(CVT) etc. works on managing frequency domain representation of an image to store data and make data detection and extraction difficult. In the context of securing financial data, it is important to consider the high security and confidentiality, which was demonstrated by Prasetyadi, G.C., Mutiara, A.B. and Refianti, R., 2017[3] research paper by encrypting data using advanced encryption standard(AES-256) through cipher block chaining(CBC) mode encryption, where in the each block of data(128 bits) is encrypted using key and Initialization vector(IV) and the outcome ciphertext XoRed with next block, hence, making it difficult to de-cipher.

de Mello, F.L. and Xexéo, J.A., 2018[20], research paper uses the machine learning to identify the strength of encryption algorithms by conducting experiments on plain text from seven different languages and seven different algorithms to encrypt and each algorithms involves ECB and CBC modes. The plaintext was encrypted and metadata was produces from the ciphertext; the metadata was analysed by six data mining algorithms. The experimentation results in ECB's mode identification rate is higher than CBC mode which means ciphertext in ECB can reveals algorithm used to encrypt the plaintext hence ECB is weaker than CBC.

Akolkar, Swati, et al., 2016[15], proposes the use of text-based steganography and visual cryptography to secure customer data while fund transfer. Swati proposed this as a technical solution where in Merchants and banks authenticate each other through OTP, which will be generated in 2 parts and transferred to user in steganographic data. OTP parts are combined and required transactions are performed. The whole objective of this proposal was to secure customer information at merchant side and prevent misuse of the data.

Moreover, Balasubramanian, and S.Geetha, 2014 [13] research mentioned confidentiality, integrity, accountability & availability as 4 security objectives to be achieved and usage of popular techniques like MFA/3D Secure SSL/TLS to safeguard data, however, as per Wei-Hsun, et al.,2018[16], these techniques are also having drawbacks as they are vulnerable to attacks e.g. POODLE (SSL 3.0/TLS 1.0) attack- CVE-2014-3566. The author reviewed prevalent techniques, as mentioned by Swati, et al.,2016[15], for data hiding and preventing theft by using visual steganography. However, for financial data like credit card/pin numbers the researchers recommend to use Highly secure credit or debit card image(HS-CDCI), where in customer uses stego card image, created by HS-CDCI software, and pin to validate the transactions between sender and receiver. Though this method of assigning software for card image enhances the security but research proposal does not recommend due to addition of complexity in handling third party software.

The author implemented the solution in 3 phases. *Phase 1,* Encrypt the customer secret data(CSD) with segmentation encryption text method(SETM) and compressed, using Huffman and RLE algorithms, before embedding . *Phase 2,* Selection of pixel by moving around CDCI by one using 8-neigbour strategy related to contrast value and hide the data. This will create a stego image and that can be shared with user. Phase 3, At receiver, deserialize the stego image and retrieve the data. This method provides highly secure and imperceptible data.

The author suggested to use Peak-signal to noise-ratio (PSNR) to measure the quality of image after embedded with secret data. Higher PSNR defines high quality of embedded image, less distortion. It Is calculated as the logarithmic ratio of the pixel value(maximum possible) to the Mean Square Error (MSE) between the original image and the stego-image. With higher % of PSNR it is difficult for hacker to identify and break security build in images using cryptography and steganography hence the approach outlined in next section is relevant to prove my research objective.

Therefore, post analysing various research papers this research proposal recommends to use AES-CBC encryption and Steganography through Transform method(DCT) using Segmentation Encryption Text Method (SETM) to secure financial data.

3. Research methodology

The below figure 1 shows overall adopted methodology to achieve security and confidentiality of financial data. The overall methodology is divided into 2 Phases which are described in detail.

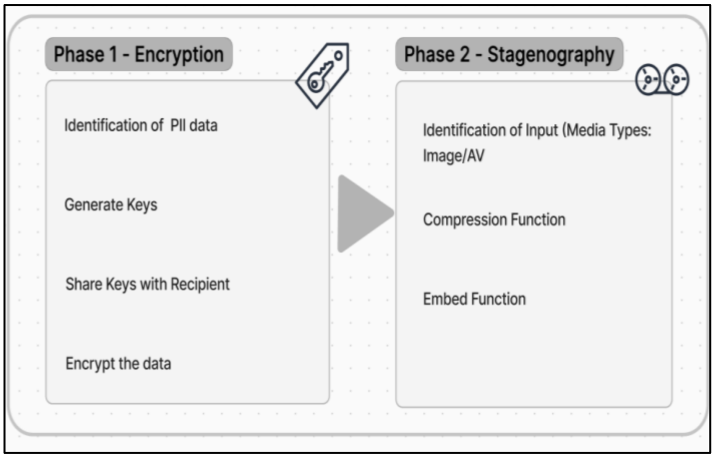


Figure 1 : Research phases

- **Phase 1 – Encryption** : This phase specifically deals with how data will be encrypted. Below are the steps mentioned.

1. Step 1: first of all we'll start with selecting the financial data to encrypt using AES-CBC. CBC encrypt plaintext in blocks and each block is XORed with previous block ciphertext and to make ciphertext unique it uses initialization vector to encrypt the plaintext.

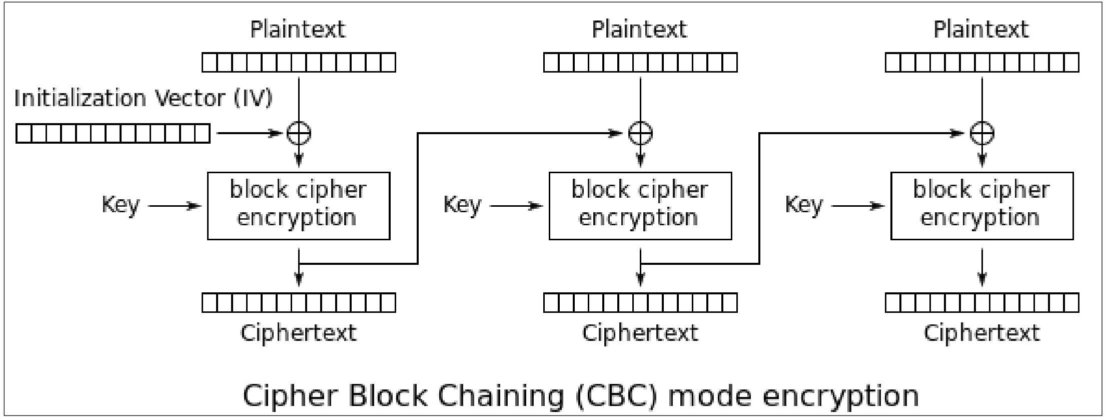


Figure 2: Cryptography using CBC

As shown in the figure , first block is Index 1, then corresponding encrypted Ciphertext defined as
$C_i = E_k (P_i \oplus C_{i-1})$
$C_0 = IV$

Similarly the ciphertext can be decrypted by reversing the process
$P_i = D_k (C_i \oplus C_{i-1})$
$C_0 = IV$

Where in, P_i *= plain text at* i^{th} *position.* C_{i-1} *= Previous(*i^{th-1}*) block Ciphertext,* C_0*= initialization vector at* 0^{th} *position.* E_k *and* D_k *are encryption and decryption functions*

- **Phase 2 – Steganography**

This phase proposes of hiding the encrypted and compressed data we generated from the source in an image using steganography technique to keep the data hidden and transfer to receive securely. It has 2 steps

1. Step 1: The secret message are of various lengths[8] and before actually going into steganography the message will be highly compressed to reduce the size to fit into image and provide more security. There are 3 main characteristics of good steganography, maximum payload, imperceptibility , and robustness. Compression enhances the payload capacity of the image validating first characteristics. This paper using DEFLATE compression[19], which is a combination of LZ77 with Huffman coding. DEFLATE is recommended as it is resilient to bit error , provide faster decompression and better support. AS per A. Gupta, A. Bansal and V. Khanduja, 2017, "Deflate provides poor compression ratio but has the fastest compression and decompression speed". For financial transaction speed is the key hence using DEFLATE. The figure shows the data is compresses using DEFLATE by referring to header 'b'x\x9c".

2. Step 2: The system insert the data over the image, which holds the secret till it reaches to receiver which will use the data using the sender's key code. The fundamental of embedding the data in an image is based on finding the key to insert the data in images and ,on the other side, the receive must use the same key. This process makes the data impregnable from attacker.

4. Design Specification

This section of the research paper discusses about the overall design of the system. Please refer (Figure 3) to below logical design specification.

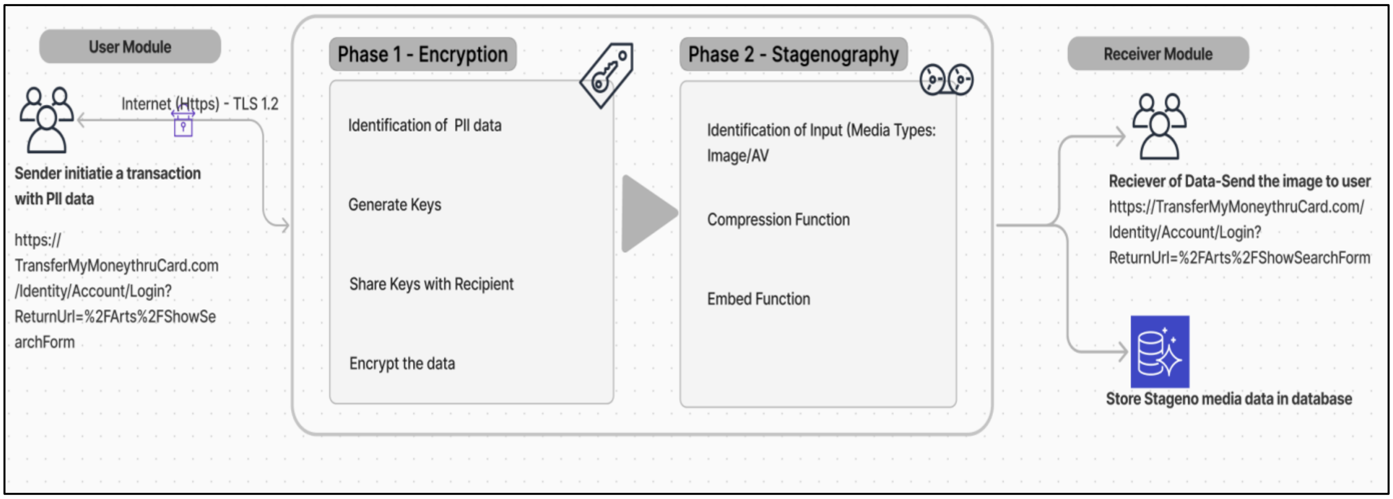

User Module
Internet (Https) - TLS 1.2
Sender initiatie a transaction with PII data
https://TransferMyMoneythruCard.com/Identity/Account/Login?ReturnUrl=%2FArts%2FShowSearchForm
Phase 1 - Encryption
Identification of PII data
Generate Keys
Share Keys with Recipient
Encrypt the data
Phase 2 - Stagenography
Identification of Input (Media Types: Image/AV
Compression Function
Embed Function
Receiver Module
Reciever of Data-Send the image to user
https://TransferMyMoneythruCard.com/Identity/Account/Login?ReturnUrl=%2FArts%2FShowSearchForm
Store Stageno media data in database

Figure 3: Logical Design of the system

The overall system consists of below 4 modules.

- User Module – where the sender or user initiate a transaction with PII data like Card Number, SSN etc. the data goes through the network layer which has encryption using TLS 1.2. The data in transit is encrypted.
- Encryption Module – Once the data pass through network layer and reaches to application layer, this module kicks in; It has further 3 functions as described below.
    - Identification of PII data – the data is either pass through GET or POST, mostly POST, this function segregates the PII vs NON-PII data.
    - Generate Keys and Share with Recipients: The public/private key pair would be created and share with the end-user so that this key would be used to decrypt. Please note: Key distribution

  - and management are not in the scope of this research paper, however, it plays an important role in this module.
  - Encrypt the data: Once the data is identified and keys are created the data is encrypted using AES-256-bit CBC encryption and passed to next module for further processing.
- Steganography Module – This module does the significant work of embedding the encrypted data in the media. For the purpose of this research the media chosen is an image. tiff format. The overall functionality is executed by 4 functions
  - Identification of input media – this function identifies the media where the data will be embedded.
  - Compression function – This function will compress the data to reduce the size of steganographic image.
  - Embed function – hide the encrypted data in the image.

- Receiver Module – The embedded image is either transferred to the receiver over https in the form of response or can be stored in the database as BLOB and can be used for further analysis.

The next section of the research paper discuss about how the logical design have been implemented.

5. Implementation

As a part of this research paper, test environment was setup which includes setup and configuration of various tools and packages.

Software/Tools setup:

1. Windows Virtual environment: As a first step, windows 11 virtual environments is setup to test various Stagno tools like Openpuff, Xiao Steganography. The objective of setting up this to do the experimentation before finalizing any tool and framework.
   Windows virtualization technology that allows to create and run virtual machines (VMs) on a Windows operating system. This enables you to run multiple operating systems on a single physical machine, providing flexibility for testing, development, and running applications in isolated environments. In this research paper, the base operating system is MacOS Sonoma(14.5). Windows 11 virtualization was setup on the top of MacOS using Parallel Virtualization desktop.
2. Python 3.01
   Python is an excellent tool for steganography and cryptography as it provides vast ecosystem of libraries used in this research paper. Libraries like PIL, Pillow, NumPy, Cryptography are used for this research paper. Python's rich libraries, ease of use, cross-platform compatibility, community support, integration capabilities, rapid prototyping, and data handling make it a powerful choice for implementing steganography and encryption solutions.
3. MetLab
   MATLAB (Matrix Laboratory) is a high-level programming language and interactive environment. It is primarily used for numerical computing, data analysis, algorithm development, and visualization. In this research paper MATLAB is widely used for image comparisons and identification of various characteristic values like PSNR,NCC, BER etc.

Configuration/Setup/Framework build: Below are the steps used to configure Python and Metlab

Step 1:

1. Encryption framework: This research paper contains Python code EncrypttheFile.py file, which encrypt the data (PII) using AES encryption/256 bit and 16Byte Initialization vector.
2. Python Cryptographic libraries are installed and used in EncrypttheFile.py.
3. The main function takes PII and Key(32bit) as an input and call *encrypt_data ()*, which uses CBC mode, IV to encode in base64.

Step 2:

4. Zlib library is installed to compress and decompress the encrypted data from Step 2. Compression.py consists of code to for that.

Step 3 :

5. The second Python file is HideDatainImage.py, consists of *Hide_data () function,* which take the image on which data to be stored and data itself. The data is output of *encrypt_data ()* function of EncrypttheFile.py

Step 4:

6. Compared the original image and steganographic image in Metlab ImageAnalysis.m file. Image processing libraries are installed to get PSNR, MSE, NCC, BER etc. values

The result of the design of this implementation will be discussed in the below section.

6. Evaluation

For the purpose of this research, I've used public images data set provided by SIPI database (University of south California). USC-SIPI [22] dataset consists of collection of images primarily to support research in image analysis and processing.

The images databases are divided in volumes with images of different types, sizes. For the purpose of this research, we have used image types .tiff and images sizes 512X512 pixels and 1024X1024 pixels. E.g. Refer to Figure 4.

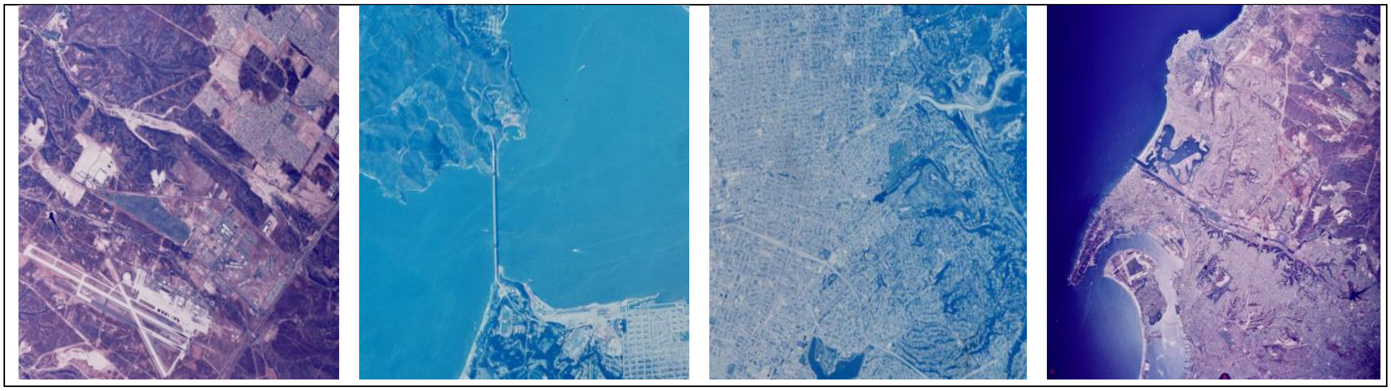
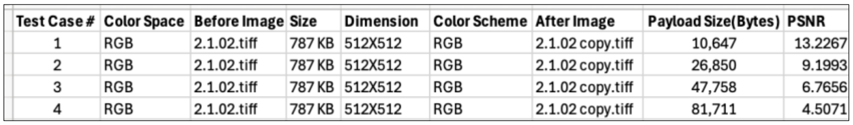


Figure 4: SIPI Dataset "Aerials "Images for testing

Steganalysis of the images can be done by two approaches. The *first approach* is objective, which depends on the values or data points of image characteristics identified before and after steganography and compare these values by applying numerical criteria using statistical tools. The *second approach* depends on human observation and judgement which sometimes leads to error hence, in this research paper we have used below characteristics.

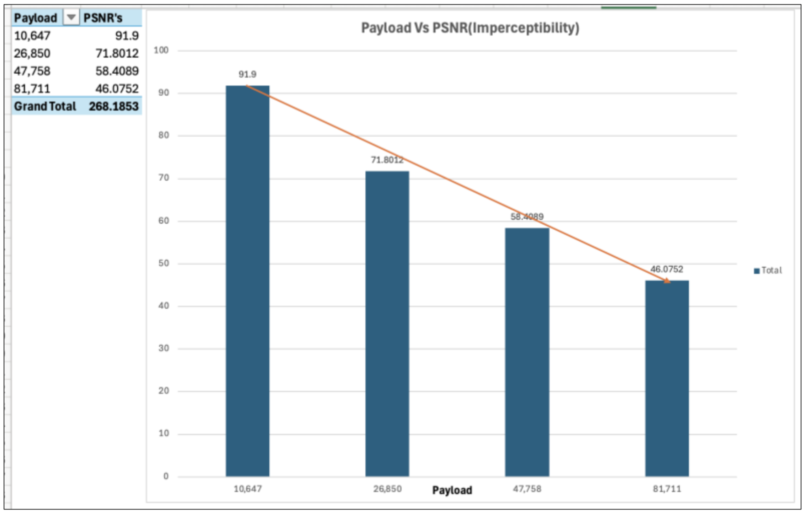
1. Peak Signal-to-Noise Ratio (PSNR): PSNR measures the ratio between the maximum possible power of a signal (in this case, the original image) and the power of the noise (the differences introduced by embedding the secret message). It is expressed in decibels (dB).
2. PSNR=10·log10(MAX^2/MSE),
    a. where MAX= Maximum possible pixel value e.g. 255 for 8 bits images
    b. MSE = squared difference between original and modified image
3. As shown in the below figure 5, e.g. Test case#1-4, as the payload size increases the PSNR value decreases.



Test Case #	Color Space	Before Image	Size	Dimension	Color Scheme	After Image	Payload Size(Bytes)	PSNR
1	RGB	2.1.02.tiff	787 KB	512X512	RGB	2.1.02 copy.tiff	10,647	13.2267
2	RGB	2.1.02.tiff	787 KB	512X512	RGB	2.1.02 copy.tiff	26,850	9.1993
3	RGB	2.1.02.tiff	787 KB	512X512	RGB	2.1.02 copy.tiff	47,758	6.7656
4	RGB	2.1.02.tiff	787 KB	512X512	RGB	2.1.02 copy.tiff	81,711	4.5071

Figure 5: Test cases executed

4. These results represent different amount of embedding, when payload increase the image get distorted a.k.a image can store data to a certain extend beyond that the image gets distorted. Our test, in figure 6, shows as payload increases from 10K to 80K the PSNR cumulative value decreases from 91.9 to 46.0752. higher PSNR is better image quality and lower means image quality is reduced.



Payload	PSNR's
10,647	91.9
26,850	71.8012
47,758	58.4089
81,711	46.0752
Grand Total	268.1853

Figure 6: Imperceptibility

5. The other characteristics was compared between images of different sizes is Mean Squared Error (MSE), it notifies greater similarity between the original image and the processed image. Low MSE indicates less error introduces during processing. In medical imaging, financial data like hiding PII data low MSE indicates preserve of image quality. MSE is very sensitive to small change in pixel value. This sensitivity is beneficial for detecting distortions introduced by processing.
6. Our test data shows, that in 24 test cases, with different image sizes, as the payload increases from 10K to 80K, the MSE also increases. Refer to figure 7.

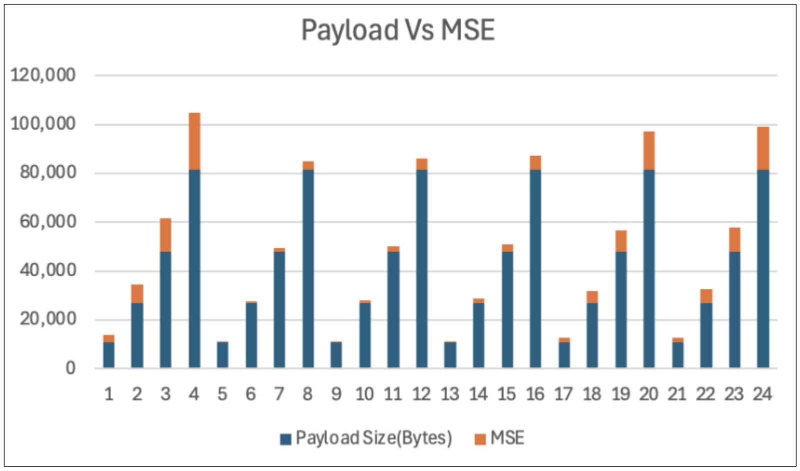
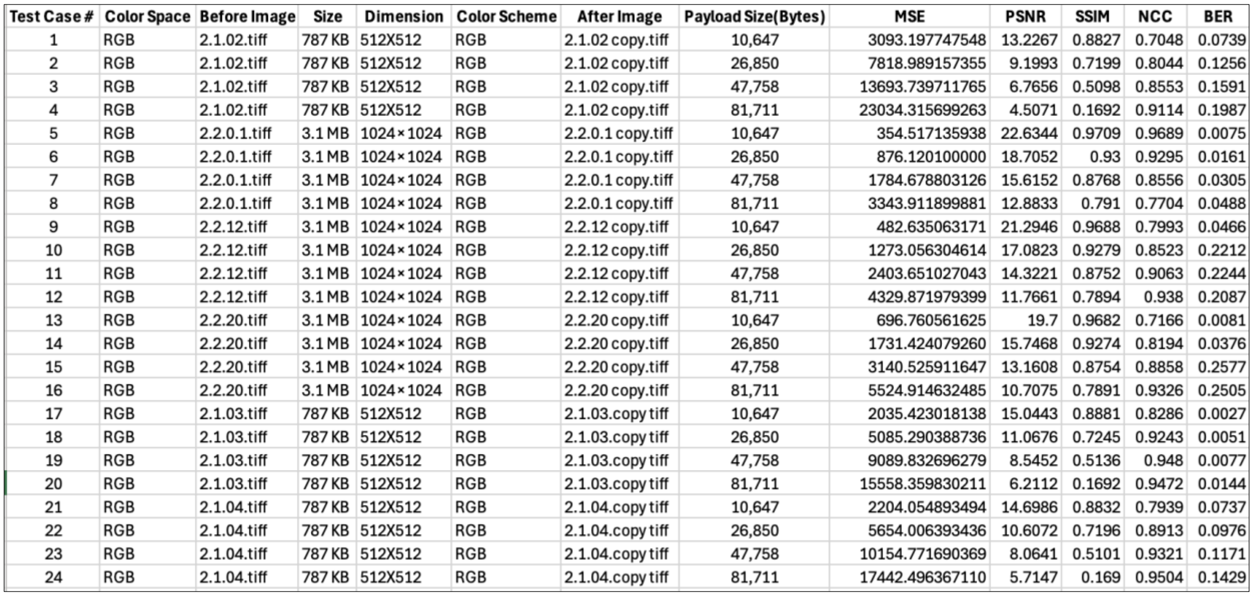


Figure 7: Increase in MSE with increase in Payload

7. Although MSE is a useful metric, it does not always correlate with perceived visual quality. Hence in our research I evaluated other characteristic like SSIM, NCC and BER to measure the robustness of the Stegano image. The below, figure 7, 24 test cases shows variation between PSNR, NCC, BER, MSE.



Test Case #	Color Space	Before Image	Size	Dimension	Color Scheme	After Image	Payload Size(Bytes)	MSE	PSNR	SSIM	NCC	BER
1	RGB	2.1.02.tiff	787 KB	512X512	RGB	2.1.02 copy.tiff	10,647	3093.197747548	13.2267	0.8827	0.7048	0.0739
2	RGB	2.1.02.tiff	787 KB	512X512	RGB	2.1.02 copy.tiff	26,850	7818.989157355	9.1993	0.7199	0.8044	0.1256
3	RGB	2.1.02.tiff	787 KB	512X512	RGB	2.1.02 copy.tiff	47,758	13693.739711765	6.7656	0.5098	0.8553	0.1591
4	RGB	2.1.02.tiff	787 KB	512X512	RGB	2.1.02 copy.tiff	81,711	23034.315699263	4.5071	0.1692	0.9114	0.1987
5	RGB	2.2.0.1.tiff	3.1 MB	1024×1024	RGB	2.2.0.1 copy.tiff	10,647	354.517135938	22.6344	0.9709	0.9689	0.0075
6	RGB	2.2.0.1.tiff	3.1 MB	1024×1024	RGB	2.2.0.1 copy.tiff	26,850	876.120100000	18.7052	0.93	0.9295	0.0161
7	RGB	2.2.0.1.tiff	3.1 MB	1024×1024	RGB	2.2.0.1 copy.tiff	47,758	1784.678803126	15.6152	0.8768	0.8556	0.0305
8	RGB	2.2.0.1.tiff	3.1 MB	1024×1024	RGB	2.2.0.1 copy.tiff	81,711	3343.911899881	12.8833	0.791	0.7704	0.0488
9	RGB	2.2.12.tiff	3.1 MB	1024×1024	RGB	2.2.12 copy.tiff	10,647	482.635063171	21.2946	0.9688	0.7993	0.0466
10	RGB	2.2.12.tiff	3.1 MB	1024×1024	RGB	2.2.12 copy.tiff	26,850	1273.056304614	17.0823	0.9279	0.8523	0.2212
11	RGB	2.2.12.tiff	3.1 MB	1024×1024	RGB	2.2.12 copy.tiff	47,758	2403.651027043	14.3221	0.8752	0.9063	0.2244
12	RGB	2.2.12.tiff	3.1 MB	1024×1024	RGB	2.2.12 copy.tiff	81,711	4329.871979399	11.7661	0.7894	0.938	0.2087
13	RGB	2.2.20.tiff	3.1 MB	1024×1024	RGB	2.2.20 copy.tiff	10,647	696.760561625	19.7	0.9682	0.7166	0.0081
14	RGB	2.2.20.tiff	3.1 MB	1024×1024	RGB	2.2.20 copy.tiff	26,850	1731.424079260	15.7468	0.9274	0.8194	0.0376
15	RGB	2.2.20.tiff	3.1 MB	1024×1024	RGB	2.2.20 copy.tiff	47,758	3140.525911647	13.1608	0.8754	0.8858	0.2577
16	RGB	2.2.20.tiff	3.1 MB	1024×1024	RGB	2.2.20 copy.tiff	81,711	5524.914632485	10.7075	0.7891	0.9326	0.2505
17	RGB	2.1.03.tiff	787 KB	512X512	RGB	2.1.03.copy tiff	10,647	2035.423018138	15.0443	0.8881	0.8286	0.0027
18	RGB	2.1.03.tiff	787 KB	512X512	RGB	2.1.03.copy tiff	26,850	5085.290388736	11.0676	0.7245	0.9243	0.0051
19	RGB	2.1.03.tiff	787 KB	512X512	RGB	2.1.03.copy tiff	47,758	9089.832696279	8.5452	0.5136	0.948	0.0077
20	RGB	2.1.03.tiff	787 KB	512X512	RGB	2.1.03.copy tiff	81,711	15558.359830211	6.2112	0.1692	0.9472	0.0144
21	RGB	2.1.04.tiff	787 KB	512X512	RGB	2.1.04.copy tiff	10,647	2204.054893494	14.6986	0.8832	0.7939	0.0737
22	RGB	2.1.04.tiff	787 KB	512X512	RGB	2.1.04.copy tiff	26,850	5654.006393436	10.6072	0.7196	0.8913	0.0976
23	RGB	2.1.04.tiff	787 KB	512X512	RGB	2.1.04.copy tiff	47,758	10154.771690369	8.0641	0.5101	0.9321	0.1171
24	RGB	2.1.04.tiff	787 KB	512X512	RGB	2.1.04.copy tiff	81,711	17442.496367110	5.7147	0.169	0.9504	0.1429

Figure 8: SSIM vs NCC vs BER

A. *Discussion*

This research paper provided an approach and implementation of hiding critical financial data using cryptography and steganography so that only authorized organization and individuals can identify and access the hidden data. As a part of experimentation using public data sets, the paper able to show that hiding the data is possible but with increase in payload the imperceptibility and size of the image increases hence reduces the acceptability of the solution by the financial organizations. For image(“2.1.02.tiff”) size 512X512 bytes RBG, Peak-Signal-to-noise ratio (PSNR) reduced by 65.9% and MSE increased to 85% in payload increases from 10K bytes to 80K bytes.

As a part of experimentation, Images of various sizes and colour coding were used. In addition to PSNR, MSE, both NCC and BER can be used to evaluate the effectiveness of embedding secret data into cover images.
For Test cases 13-16 and Test cases 17-20 the NCC was increased from 0.7166 to 0.9326 and 0.8286 to 0.9472 respectively in the case when payload increase from 10K to 80K bytes. High NCC values indicate that the cover image remains visually similar after data embedding, while low BER values indicate that the data can be retrieved accurately without significant errors, in test cases from 1-24 there is a little movement in BER, which can be seen in the below figure 9.

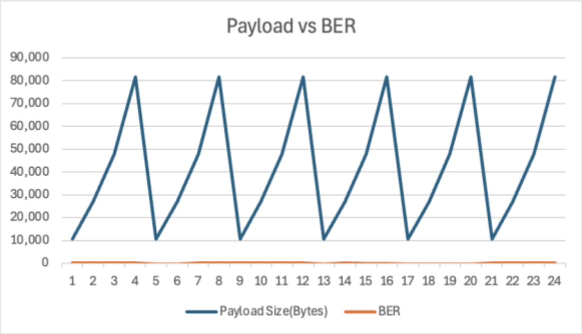


Figure 9: BER values for 24 test cases with increase in payload

Overall, the experimentation was structure is phases, encryption, steganography and data analysis of before and after picture. The test results show that impact on the parameters with payload increase. **Please refer to "Thesis – Data Analysis Sheet for more details"**

Parameters	**Payload increase from 10K to 80K bytes**
PSNR	Decreases ↓
MSE	Increases ↑
NCC	Increases ↑
BER	Slight increase ↑

Hence, to increase the efficacy and efficiency we need to be very cognizant on how much data we need to encrypt and embed through steganography.

I believe that approach mentioned in this research paper holds significant potential in the financial market for concealing customer data. For instance, it can be used to hide information within credit cards, virtual cards, or QR codes. However, with the advancement of GPU technologies and advanced hardware capabilities, in future, the strength of AES encryption (256/512 bit) will get data challenged hence both the techniques should be used together.

Based upon the work done so far in this research paper, I believe approach given in the paper is substantial moreover, this work can be extended further to other areas of technologies like secure communication, database security, digital water marking etc. even some of the good work done by earlier researcher Swati, et al., 2016[15] & Kadhim, Inas Jawad, et al.2019 [23] in the field of secure communications and Payments.

7. Conclusion and Future Work

The objective of this paper is to research is not only to encapsulate the financial data but also to prove that the whole process works. The implementation proposal for the research was successfully completed. This project can be enhanced to better support all images & video/audio types. Utilizing tools such as Metlab, Python and available datasets can facilitate this improvement. It is advisable to consider this approach and highly recommended to pursue it in future work. For real-time applications of this proposal, there are areas like Confidential communication, Copyright protection, Database systems, Access controls, Secure data transmission, where the mentioned approach can be used.

REFERENCES

[1] Divya, A., and S. Thenmozhi. "Steganography: Various Techniques In Spatial and Transform Domain." *Int. J. Adv. Sci. Res. Manag* 1.3 (2016): 8189

[2] Wu, H., Liu, X. and Tang, W., 2011, June. A fast GPU-based implementation for MD5 hash reverse. In 2011 IEEE International Conference on Anti-Counterfeiting, Security and Identification (pp. 13-16). IEEE.

[3] Prasetyadi, G.C., Mutiara, A.B. and Refianti, R., 2017, November. File encryption and hiding application based on advanced encryption standard (AES) and append insertion steganography method. In 2017 Second International Conference on Informatics and Computing (ICIC) (pp. 1-5). IEEE.

[4] Sharma Rohini, Anil Kumar" A Secure Image Steganography Based on RSA Algorithm and Hash-LSB Technique" International Journal of Advanced Research in Computer Science and Software Engineering Volume 3, Issue 7, July 2013

[5] Halder, R., Sengupta, S., Ghosh, S. and Kundu, D., 2016. A secure image steganography based on rsa algorithm and hash-lsb technique. IOSR Journal of Computer Engineering (IOSR-JCE), 18(1), pp.39-43.

[6] George, Thomas Ronil (2018) Securing password hashes from SQL injection attacks using Image Steganography. Masters thesis, Dublin, National College of Ireland.

[7] Sarmah, Dipti Kapoor ; Bajpai, Neha. / Proposed System for data hiding using Cryptography and Steganography. In: International Journal of Computer Applications . 2010 ; Vol. 8, No. 9. pp. 7-10.

[8] Shafry, Mohd & Rahim, Mohd & Falah, Yusur & Hashim, Mohammed & Zainal, Anazida. (2020). Hiding Financial Data In Bank Card Image Using Contrast Level Value And Text Encryption For Worthiness A Robust Steganography Method. 27. 2783-2801.

[9] PSNR of Image in MetLab, https://stackoverflow.com/questions/40395657/psnr-of-image-in-matlab

[10] Murakami, T., Kasahara, R. and Saito, T. (2010). An implementation and its evaluation of password cracking tool parallelized on gpgpu, Communications and Information Technologies (ISCIT), 2010 International Symposium on, IEEE, pp. 534–538.

[11] Shannon, C.E., 1949. Communication theory of secrecy systems. The Bell system technical journal, 28(4), pp.656-715.

[12] Wang, H & Wang, S, "Cyber warfare: Steganography vs. Steganalysis", Communications of the ACM, 47:10, October 2004

[13] Chelliah, Balasubramanian, and S. Geetha. "Enhancing E-Payment Security through Biometric Based Personal Authentication Using Steganography Scheme–B-PASS." *International Conference on Security in Computer Networks and Distributed Systems*. Springer, Berlin, Heidelberg, 2014.

[14] Oruganti, Madhu, et al. "Difference Expansion based Near Reversible Data Hiding Scheme." *First International Conference on Artificial Intelligence and Cognitive Computing*. Springer, Singapore, 2019.

[15] Akolkar, Swati, et al. "Secure Payment System using Steganography and Visual Cryptography." *International Journal of Computing and Technology* 3.1 (2016).

[16] Lee, Wei-Hsun, et al. "A peer-to-peer transaction authentication platform for mobile commerce with semi-offline architecture." *Electronic Commerce Research* 18.2 (2018): 413-431.

[17] Morkel, T., Eloff, J. H. and Olivier, M. S. (2005). An overview of image steganography., ISSA, pp. 1–11

[18] Splunk, "The CISO Report 2024"

[19] A. Gupta, A. Bansal and V. Khanduja, "Modern lossless compression techniques: Review, comparison and analysis," 2017 Second International Conference on Electrical, Computer and Communication Technologies (ICECCT), Coimbatore, India, 2017, pp. 1-8, doi: 10.1109/ICECCT.2017.8117850.

[20] de Mello, F.L. and Xexéo, J.A., 2018. Identifying Encryption Algorithms in ECB and CBC Modes Using

[0]

[21] NASDAQ, Verafin, 2024, Global Fiancial Crime Report

Computational Intelligence. *J. Univers. Comput. Sci.*, *24*(1), pp.25-42.

[22] https://sipi.usc.edu/database/ "Public Image database from University of Southern California"

[23] Kadhim, Inas Jawad, et al. "Comprehensive survey of image steganography: Techniques, Evaluations, and trends in future research." *Neurocomputing* 335 (2019): 299-326.