# DESIGN AND IMPLEMENTATION OF HIGH ACCURATE AND ERROR FREE MATRIX MEMORY USING PRPG

[1]THATI. CHANDANA, [2]Y. BHASKARA RAO
[1]M.Tech Scholar, Dept of ECE, Malineni Lakshmaiah Womens Engineering College, A.P, India
[2]Assistant Professor, Dept of ECE, Malineni Lakshmaiah Womens Engineering College, A.P, India

**ABSTRACT: In this paper design and implementation of high accurate and error free matrix memory using PRPG is implemented. The main intent of this project is to reduce the delay and improve the performance of system. Initially input address and input data is given. these input data performs decoding operation using address decoder. If there are any errors in obtained data PRPG will detect and correct it and gives the accurate data. Address control unit decodes address of data in two ways they are row decoder and column decoder. Row decoder decodes the data in row format and column decoder will decodes the data in column format. At last the row and column data will be saved in memory matrix. From this data will perform read and write operations. This is simulated using Xilinx technology. From simulation results it can observe that effective output is obtained in terms of delay and area.**

**KEY WORDS: Error Detection and Correction Codes (EDC), RAM (Random Access Memory), Address Generator, Encoder and Decoder, PRPG.**

## I.INTRODUCTION

In recent years error-correcting codes (ECCs) have been used increasingly to enhance the system reliability and the data integrity of computer semiconductor memory subsystems. As the trend in semiconductor memory design continues toward higher chip density and larger storage capacity, ECCs are becoming a more cost-effective means of maintaining a high level of system reliability.

A memory system can be made fault tolerant with the application of an error-correcting code; i.e., the mean time between "failures" of a properly designed memory system can be significantly increased with ECC. In this context, a system "fails" only when the errors exceed the error-correcting capability of the code. Also, in order to optimize data integrity, the ECC should have the capability of detecting the most likely of the errors that are uncorrectable.

Transient errors and permanent faults in memory chips are well known reliability issues in computer systems.
Error Detection and Correction (EDAC) codes, also called Error-Correcting Codes (ECCs) are the prevailing solution to this problem. Typically, to accommodate extra bits the memory bus architecture is extended and to detect and correct errors the coding and checking circuitry is added. But due to cost considerations this additional hardware can be sometimes removed. Hardware redundancy techniques, such as duplication or Triple Modular Redundancy (TMR), can be one solution, but they are very expensive.

Bit-flips caused by Single Event Upsets (SEUs) are a major problem in memory chips and use of EDAC codes remained an effective solution to this problem. To implement these codes on hardware, extra memory chips and encoding/decoding circuitry is required. In systems where EDAC hardware is not available, but when the hardware support is not available then the protection is provided through software. Software implemented EDAC could be a better choice than hardware EDAC, because it can be used with a simple memory system and it provides the flexibility of implementing more complex coding schemes.

Single event upsets (SEUs), power fluctuations or electromagnetic interference

can be the reason for soft errors. As process technology scales down to small nano metres, high-density, low cost, high performance integrated circuits, characterized by high operating frequencies, low voltage levels and small noise margins will be increasingly susceptible to temporary faults. Moreover, single-event upsets (SEUs) and single-event transients (SETs) generated by atmospheric neutrons and alpha particles severely impact field-level product reliability, not only for memories, but also for logics in very deep sub-micron technologies. When these particles hit the silicon bulk, they create minority carriers which if collected by the source/drain diffusions, could change the voltage level of a node.

The number of errors generated in the failure of a memory chip is largely dependent on the chip failure type. For example, a cell failure may cause one error, while a line failure or a total chip failure in general causes more than one error. For ECC applications, the memory array chips are usually organized so that the errors generated in a chip failure can be corrected by the ECC. In the case of SEC codes, the one-bit-per-chip organization is the most effective design. In this organization, each bit of a codeword is stored in a different chip; thus, any type of failure in a chip can corrupt, at most, one bit of the codeword. As long as the errors do not line up in the same codeword, multiple errors in the memory are correctable.

Storage spaces which can store data are called memories. In memory, data is encoded while storing and decoded while retrieving. By using separate storage elements, making them as cluster information can be stored. Each memory cell is capable to store a bit of data. While storing the data, soft errors which occur due to particle radiation are a huge threat for the reliability of memories. When radiation particle touches the sensitive area of the system, errors will occur. Due to which the data stored in memory is corrupted.

Error-correcting codes used in early computer memory systems were of the class of single-error-correcting code and double error-detecting (SEC-DED) codes invented by R. W. Hamming. A SEC-DED code is capable of correcting one error and detecting two errors in a code word. The double-error detecting codes have capability to serve as guard against data loss. In 1970, a new class of SEC-DED codes called odd-weight-column codes was published by Hsiao. With the same coding efficiency, the odd-weight-column codes provide improvements over the Hamming codes in speed, cost and reliability of the decoding logic.

The errors in semiconductor memories can be basically divided into two types: hard errors and soft errors. A hard error occurs when a memory location of hardware becomes permanently defective. It is an irreversible error caused by connection failures like internally shorted or open leads. Soft errors are temporary and random in time and locations. They may occur during one particular memory cycle time but disappear in the next cycle. The soft errors result from system noise, power surges, atmospheric interference and alpha-particle radiations.

High-density, low-voltage levels, small feature size and small noise margins make the memory chips increasingly susceptible to faults or soft errors. Errors introduced due to the external radiation or electrical noise rather than the design or manufacturing defects are known as soft errors. They are caused by high energy neutrons and alpha particles hitting the silicon bulk resulting in

the production of large number of electron-hole pairs. The accumulated charge may be sufficient to flip the value stored in a cell thus causing bit inversion, resulting in soft error. Hence the effects of radiation are bit-flips occurring in the information stored in memory elements. Due to the relentless shrinkage in the device dimensions, the particles that were once considered negligible are now proving to be significant enough to cause upsets. Such errors are identified as soft errors since, although they corrupt the value stored in the cell, they do not permanently damage the hardware.

## II. LITERATURE SURVEY

**A Robust Code for MBU Correction till 5-Bit Error [1]**

Memory is a component, playing a significant role in electronic systems. The use of static random access memories (SRAMs) is increasing in multimedia and system on chip applications. The key challenge faced by SRAMs is soft errors induced by the radiation that cause the change of values in memory cells. Error correction codes (ECCs) are used to face this challenge. As the technology is scaling down the chances of multiple bit upsets (MBUs) is increasing. So ECCs with higher correction ability is needed. Many ECCs have been proposed to face the challenge of MBUs.

**Error Detection and Correction in Semiconductor Memories using 3D Parity Check Code with Hamming Code [2]**

Data stored in memory or buffer needs Error Detection and Correction (EDAC). Errors occur due to supply voltage fluctuations and/or noise due to electromagnetic interference or external radiation. These errors could be either temporary or permanent. In this paper, an EDAC method is proposed to detect and correct errors based on 3D parity check. In the encoder, the data bits are arranged in a matrix format and then parity bits are calculated for each row, column and diagonal. Errors present in parity bits are detected and corrected using Hamming code. Regeneration of data bits and Syndrome calculation at the decoder helps in detecting and correcting the error bits in the data. The 3D Parity check code can correct up to 3 bits of any combination of errors in the data and the Hamming code can correct up to 3 bits in the parity, if they occur in specific combinations. Thus, this method can detect and correct errors in both data and parity bits. This method achieves higher reliability by having a slight tradeoff in area and power consumption compared to other similar methods.

**Low Redundancy Matrix-Based codes for Adjacent Error Correction with Parity [3]**

As CMOS technology scales down, multiple cell upsets (MCUs) caused by a single radiation particle have become one of the most challenging reliability issues for memories in space applications. In general, bits affected by MCUs are usually physically close. Error correction codes (ECCs) are commonly used to protect memory against MCUs. Recently, Matrix-based codes are an interesting option due to their low complexity decoding. The number of parity bits matrixbased codes required, which is related to the redundancy cells in memories, is not small. In this paper, a low redundancy scheme for matrix-based codes is presented. Based on a new matrix arrangement, the proposed scheme combines the extended Hamming codes per row and parity codes per column with parity sharing.

**Horizontal-Vertical Parity and Diagonal Hamming Based Soft Error Detection and Correction for Memories [4]**

External radiations create soft errors which are turning into an undeniable critical issue. Customarily, Single Error Correction (SEC)

code which can detect and correct 1-bit error per memory word is used to rectify soft errors. As errors turn out to be more common, the SEC methodology becomes inefficient. In this paper, Horizontal-Vertical Parity and Diagonal Hamming (HVPDH) method is proposed for detection of up to 8-bit errors and correction of 1-bit error, all combinations of 2-bit errors and most combinations of 3, 4 and 5-bit errors for memories. The aim here is to incorporate an encoder and decoder which will be effective in detecting and correcting errors. The encoder and decoder use three parity sets, namely horizontal, vertical and grouped diagonal Hamming parities. As per the analysis, higher code rate is achieved by using HVPDH method in memories when compared to the existing methods.

**Matrix code based Multiple Error Correction Technique for N-bit memory data [5]**

Constant shrinkage in the device dimensions has resulted in very dense memory cells. The probability of occurrence of multiple bit errors is much higher in very dense memory cells. Conventional Error Correcting Codes (ECC) cannot correct multiple errors in memories even though many of these are capable of detecting multiple errors. This paper presents a novel decoding algorithm to detect and correct multiple errors in memory based on Matrix Codes. The algorithm used is such that it can correct a maximum of eleven errors in a 32-bit data and a maximum of nine errors in a 16-bit data. The proposed method can be used to improve the memory yield in presence of multiple-bit upsets. It can be applied for correcting burst errors wherein, a continuous sequence of data bits are affected when high energetic particles from external radiation strike memory, and cause soft errors. The proposed technique performs better than the previously known technique of error detection and correction using Matrix Codes.

## III. PROPOSED SYSTEM

The below figure (1) shows the block diagram of proposed system. Initially input address and input data is given. These input data performs decoding operation using address decoder. If there are any errors in obtained data PRPG will detect and correct it and gives the accurate data. Address control unit decodes address of data in two ways they are row decoder and column decoder. Row decoder decodes the data in row format and column decoder will decodes the data in column format. At last the row and column data will be saved in memory matrix. From this data will perform read and write operations. This is simulated using Xilinx technology

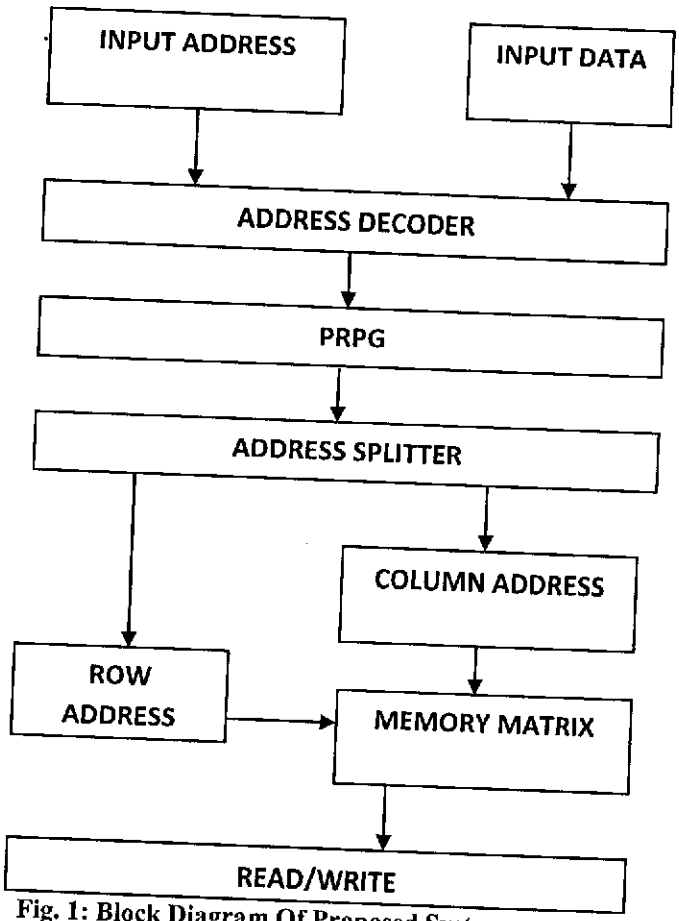



Fig. 1: Block Diagram Of Proposed System

**Algorithm:**

**Step-1:** Initially input address and input data is given.

**Step-2:** These input data performs decoding operation using address decoder.

**Step-3:** If there are any errors in obtained data PRPG will detect and correct it and gives the accurate data.

**Step-4:** Address control unit decodes address of data in two ways they are row decoder and column decoder.

.

**Step-5:** Row decoder decodes the data in row format and column decoder will decodes the data in column format.

.

**Step-6:** At last the row and column data will be saved in memory matrix.

.

**Step-7:** From this data will perform read and write operations.

## IV. RESULTS AND DISCUSSION

The below figure (2) shows the RTL schematic of proposed system.

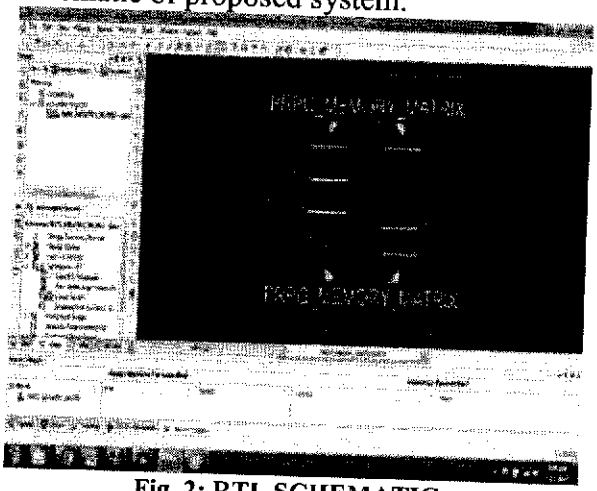


**Fig. 2: RTL SCHEMATIC**

The below figure (3) shows the technology schematic of proposed system.

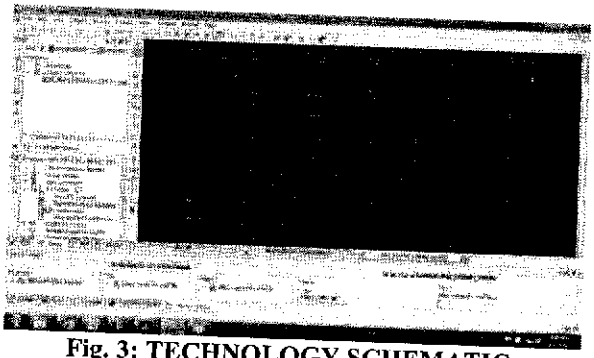


**Fig. 3: TECHNOLOGY SCHEMATIC**

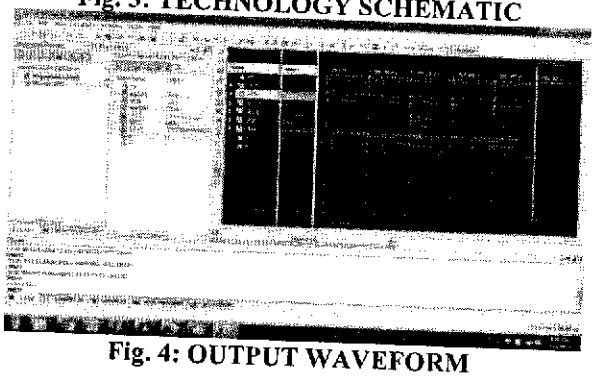


**Fig. 4: OUTPUT WAVEFORM**

## IV. CONCLUSION

Design and implementation of high accurate and error free matrix memory using PRPG was implemented. Address control unit decodes address of data in two ways they are row decoder and column decoder. From this data will perform read and write operations. This is simulated using Xilinx technology. From simulation results it can observe that effective output is obtained in terms of delay and area.

## V. REFERENCES

[1] K. Sai Karan, N. Srikanth, Sonali Agrawal,"Robust Code for MBU Correction Till 5-Bit Error," International Conference on Communication and Electronics Systems, 2019.

[2] Shivani Tambatkar, Siddharth Narayana Menon, Sudarshan.V, M.Vinodhini and N.S.Murty , "Error Detection and Correction in Semiconductor Memories using 3D Parity Check Code with Hamming Code ," International Conference on Communication and Signal Processing, April 6-8, 2017, India

[3] Shanshan Liu, Liyi Xiao, Jie Li, Yihan Zhou, and Zhgang Mao, "Low Redundancy Matrix-Based codes for Adjacent Error Correction with Parity ," 18th International Symposium on Quality Electronic Design (ISQED),2017.

[4] Paromita Raha, M Vinodhini, N. S. Murty, "Horizontal-Vertical Parity and Diagonal Hamming Based Soft Error Detection and Correction for Memories," Jan. 05 – 07, 2017, Coimbatore, INDIA.

[5] Sunita M.S and Kanchana Bhaaskaran , " Matrix code based Multiple Error Correction Technique for N-bit memory data" , International Journal of VLSI design Communication Systems (VLSICS) Vol.4, No.1, February. 2013.

[6] N. N. Mahatme, B. L. Bhuva, Y. P. Fang, and A. S. Oates, "Impact of strained-Si PMOS transistors on SRAM soft error rates," IEEE Trans. Nucl. Sci., Vol. 59, No. 4, pp. 845–850, August. 2012.

[7] E. Ibe, H. Taniguchi, Y. Yahagi, K. Shimbo and T. Toba, "Impact of scaling on neutron-induced soft error rate in SRAMs From a 250 nm to a 22 nm Design Rule", IEEE Trans. on Electron Devices, Vol. 57 , No. 7, pp. 1527-1538, July. 2010.

[8] Naeimi H, Dehon A. "Fault secure encoder and decoder for nanomemory applications". IEEE Trans Very Large Scale Integr (VLSI) Syst , Vol. 17, No. 4 pp. 473-486, 2009.

[9] D. Radaelli, H. Puchner, S. Wong, and S. Daniel, "Investigation of multi bit upsets in a 150 nm technology SRAM device," IEEE Trans. Nucl. Sci., Vol. 52, No. 6, pp. 2433–2437, Dec. 2005.

[10] R.C. Baumann, "Radiation-induced soft errors in advanced semiconductor technologies", IEEE Trans Dev Mater Reliab., Vol.5, No.3, pp. 301-316, Sep. 2005.

[11] C. L. Chen and M. Y. Hsiao, "Error-correcting codes for semiconductor memory applications: A state-of-the-art review," IBM J. Res. Develop. , Vol. 28, No. 2, pp. 124–134, Mar. 1984.

[12] Fernanda Lima, Luigi Carro, Ricardo Reis "Designing Fault Tolerant Systems into SRAM-based FPGAs" Anaheim, California, USA, DAC'03, June 2-6, 2003.

ZKG INTERNATIONAL

ISSN: 2366-1313

# Internet Of Things Based Student Monitoring Mechanism Using RFID

[1]L. Jaya Adi Lakshmi, [2]A.Roja Venkata Siva Saikumari,[3]M. Harika, [4]Gattamaneni Manisha,[5]Dr. K. Gouthami

[1]BTech Student, Dept.of ECE, Malineni Lakshmaiah Womens Engineering College, Guntur. AP

[2]BTech Student, Dept.of ECE, Malineni Lakshmaiah Womens Engineering College, Guntur, AP

[3]BTech Student, Dept.of ECE, Malineni Lakshmaiah Womens Engineering College, Guntur, AP

[4]BTech Student, Dept.of ECE, Malineni Lakshmaiah Womens Engineering College, Guntur, AP

[5]Professor (Ph.D), Dept.of ECE, Malineni Lakshmaiah Womens Engineering College, Guntur, AP

***Abstract**: RFID is a nascent era, deeply rooted in its early trends in using Radar 1 as a weapon for enemy aircraft sometime in World War II. A large number of industries have taken advantage of the RFID era's advantages to improve sectors such as the military, sports activities, security, airlines, animal farms, healthcare, and other areas. In this proposed system, law students receive an RFID tag. Therefore, the stats saved on this card are called the identity/presence of the individual. Once the requester places the card in front of the RFID card reader, it reads the records and checks them against the stats stored inside the 8051-family microcontroller. If the stats match, it displays a message on the LCD screen to confirm identification. Input from this student; Otherwise, she delivers a message refusing to help. The student's attendance status can be retrieved from this system by pressing the Reputation button attached to the microcontroller. Thus, a lot of time is saved as all students' attendance is stored directly in the database.*

***Keywords:** RFID, Internet of Things Based attendance system, Microcontroller.*

## I. INTRODUCTION

The Internet of Things (IoT) concept has attracted increasing interest these days from both academic and business circles. The Internet of Things is where devices (even animals or people) have accurate identifiers and the ability to automatically transmit records through a community without requiring human-computer interaction [1]. The Internet of Things is a situation in which devices (even animals or humans) are equipped with accurate identifiers and the ability to automatically transmit data over a network without requiring human-computer interaction. RFID forms a fundamental building block of IoT, where RFID devices are micro-wi-fi chips that are used to identify devices by computer. Student attendance is a vital part

ISSN: 2366-1313

of daily instruction. Traditionally, teachers are given the task of calling out class names. Thus, this is time-consuming, and you also no longer have the flexibility to produce reports or logs. Researchers have proposed several technologies to get rid of the guided attendance procedure of signatures on papers, including barcode-based attendance systems, facial recognition, and fingerprint identification. However, these structures suffer from some obstacles and difficulties [2]. The most common method of tracking student attendance systems is by manually taking a list of names or recording them in the attendance sheet. For a classroom full of energy, both strategies are stressful. The name roll method quickly risks false attendance in a large class and takes longer to say the names of all the college students. Significant problems also arise with converting paper facts into an electronic form for students' electronic records to calculate total attendance in several grades (e.g., subject, exam program, school, or university). In addition to the above risks, the most significant common drawback is that these strategies require more equipment. A proposed device has been developed to address these risks. The main benefits of the proposed system are flexible use, no device costs, no time loss, and easy access [3]. Classroom attendance machine is mainly based on face reputation generation, combined with the era of RFID. It effectively emphasized the identification of schoolchildren within the class. Real-time testing of the rule set fully meets the time attendance requirements within elegance, reduces the value of attendance in the classroom, and successfully solves signature troubles and other problems. For Internet server structures, the XAMPP program is used. XAMPP is the software that contains the complete Internet optimization environments for PHP, Apache, and MySQL. XAMPP is an accessible, open-source web server for optimizing web-based packages natively. SQL is a single-reason programming language that manipulates records in a relational database control system. The MySQL device in XAMPP is PHPMyAdmin. To maintain a real identity on the student ticket, MySQL is needed. In MySQL, four tables were created, including the employee table, the student table, the student attendance desk, and the student grades.

## RFID TECHNOLOGY

RFID is popular in Radio Frequency Identification; the current idea of the Internet of Things (IoT) era is similar to barcode devices but with a slightly more advanced concept. It works by using the mobile signal and receiving it through the antenna and the integrated circuit. It

ISSN: 2366-1313

contains components that are RFID tag and an RFID reader.

**RFID Tag**

An RFID tag is a digital tag that exchanges data with an RFID reader via radio waves. Almost all RFID tags contain special elements, an antenna, and an integrated circuit (IC). The antenna is used to receive radio frequency waves, and the IC used to process and record recordings.

**RFID Reader**

The RFID reader is a device to obtain records of RFID tags used for male or female music. RFID uses radio waves to transmit records from the tag to the reader.

**WORKING MECHANISM**

The RFID tag receives the signal from the reader through the antenna and is charged. The loaded tag then sends the feedback back to the antenna. The antenna reads the data and sends it to the reader. Finally, the reader reads and follows the necessary commands.

## II. LITERATURE SURVEY

Kariapper et al. [4] This paper discusses the latest development and application of IoT attendance systems using RFID. This analysis found that RFID ushers in a completely new era of computerized assistive devices and provide much better accuracy and performance than traditional paper-based systems. The hybrid version is necessary and must achieve greater security, reliability, lower price, and better device performance. This paper will use the RFID 4 Biometric frame machine with a sufficient number of RFID tag readers and fingerprint devices to verify the device's security and reliability. The web server should be replaced with PHP and MYSQL, which are cheaper, with better accuracy and overall performance overkill. This hybrid model can be applied and practiced efficiently in schools and better training institutions.

Madhu et al. [5] Nowadays, we have witnessed a sudden surge in the use of Radio Frequency Identification (RFID) systems in business technologies, fitness, agriculture, transportation, etc. Moreover, the Internet of Things is thriving in parallel. So, by using these, an effort has been made to solve attendance tracking and control problems. Attendance Management System is an IoT application with Raspberry Pi 3 and RFID technology to reduce feeding time with traditional daily attendance systems in colleges and institutions. Therefore, everything here, in turn, is automated. An effort has also been made to increase the android software (application) and help scholars to watch their presence everywhere and at any time.

ISSN: 2366-1313

B.M Sri Madhu et al. [6] Recently, we have seen a sudden increase in the use of Radio Frequency Identification (RFID) systems in business technology, healthcare, agriculture, transportation, etc. Moreover, the Internet of Things is thriving in parallel. So, by using these, an attempt is made to solve the problems of attendance tracking and control. Attendance Management System is an IoT application with Raspberry Pi 3 and RFID technology, a way to reduce feeding time with the help of traditional daily attendance devices in schools and organizations. Therefore, everything here, in turn, is automated. An attempt has also been made to develop an Android app (application) and help students to see their presence everywhere and at all times.

Turkane et al. [7] The online support method is helpful for workers who perform activities outside the workplace or people with multiple agendas. Recently, we have seen a sharp increase in the use of face-detection focus structures in business technology, fitness, agriculture, transportation, and many more. Moreover, the Internet of Things is thriving in parallel. Therefore, by using them, an effort has been made to solve the problems of attendance tracking and control. Online Attendance Management System is the implementation of the Internet of Things through the Raspberry Pi and face detection approach to reduce the time it takes for a traditional device to record daily attendance in schools and organizations. This system describes assistance without human intervention. In this strategy, the camera is constantly inside the meeting room and will take the picture; the faces are detected, then it is diagnosed with the database, and at some point, the attendance is determined. So, everything here, in turn, becomes automated. An attempt has also been made to develop an Android app(s) and help students see their presence everywhere and at every time.

Mathew Turk et.al, [8hHere they built a near real-time computing device that could find and adjust a person's head, then stop the man or woman by comparing their facial features to those of known humans. The computational methodology adopted in this framework is motivated by every body structure and the idea of facts, as well as by the practical requirements of general performance and accuracy near real-time. This method treats the face popularity problem as a typical 2D reputation problem rather than a 3D geometry retrieval order, exploiting that such appearances are often rectilinear and thus likely to be represented by the tangible association of 2D brands. His

ISSN: 2366-1313

investigations show that the eigenfaces method can work very accurately, albeit with enormous "hard-to-grasp" separation charges, and is, therefore, probably suitable for these beams. The future scale of this project has become, in addition to facial recognition, the utilization of the investigation of faces to determine the sexual orientation of the subject and the interpretation of facial expressions.

## III. PROPOSED METHODOLOGY

### Internet of Things (IoT)

The Internet of Things (IoT) is an intercom or network of several devices on the body, such as cars and apartments, integrated with sensors, software, electronic hardware, and connectivity that helps retrieve and alter records. It allows the discovery and management of objects through the available network infrastructure, integrating the physical environment and its elements with portable systems. It provides superior connectivity between devices and systems passing through device-to-device courtship.

The project method is that once the person entering has been read by PIR sensor 1, the RFID reader will be activated and receive only one card at a time until the other PIR sensor detects the character. This way, until the PIR 2 sensor detects that the character is moving in the class, attendance will not be updated in the database. Here we used XAMPP to create the database in a PHP script for the Apache server. Attendance will be updated directly on this web page, and we have also designed a basic app (application) for Android through which students can check their attendance directly from the app on their mobile phones.

The proposed system has been explained with the help offollowing steps:

**Step 1:** Start the RFID Reader

**Step 2:** Initialize the LCD Screen

**Step 3:** Initialize UART (Universal AsynchronousReceiver/ Transmitter)

**Step 4**: Send scanned UID of RFID card data toRaspberry Pi Model

**Step 5:** Search and match the UID and extract therelevant student information

**Step 6:** Compare detected student id, date and timewith class time table and if match found then markthe presence.

We have widely used passive infrared sensors to solve the problem of a support agent. The first PIR sensor will first detect the movement of the man or woman by sensing their body heat and will give an output of 1. Once the PIR output is 1, the RFID reader is programmed so that the requester does not skip the second PIR,

ISSN: 2366-1313

and the RFID reader can study only one RFID tag. Now the student will only touch her card, not the one not there. Then the student will enter the lecture room, and the second PIR sensor will read out loud. Once the second PIR is high, the student can mark attendance for that single issue, and the growth will be calculated by 1. Similarly, if the second PIR reads first, then the first, it counts will go down through one, and the school will find out there is an agent because the numbers will vary. It can update Raspberry Pi records in the teacher's database without delay, and they don't want to do the hard work of getting benefits support. All teachers can have a username and password to log into the database. We used XAMPP (X: Cross-Platform (WINDOWS, LINUX, MAC OS) A: Apache M: My SQL P: PHP P: PEARL) to build the database on time.

We have also developed an ANDROID APP so that college students can view their daily attendance on their mobile phones, making it easier for you to take your regular attendance test. This android app contains details about the student, his call, USN, tracks recorded, group of classes attended, list and number of instructions taken, and eligibility reputation, which makes it easy for the student to keep a song for help reputation and thus live consciously. This is an open platform, and any man or woman can look into it by simply entering the student's name and the position whose attendance you wish to understand.

**Arduino Uno**

Arduino Uno is an open-source microcontroller board mainly based on the ATmega328P microcontroller developed by Arduino. The board is equipped with virtual and analog input/output (I/O) modules that can be connected to various expansion boards (nine displays) and other circuits. The board has 14 digital I/O pins (six suitable for PWM output) and six analog I/O pins and is programmable using the Arduino IDE (Integrated Development Environment) via a Type B USE cable. It can be powered through a USB cable or an external 9V battery, although it accepts voltages between 7 and 20V. It is also like Arduino Nano and Leonardo. Hardware reference design submitted under the Creative Common website. ShareAlike Attribution 2.5 AM License.

**EM-18 RFID reader**

The EM-18 RFID reader module operating at 125 kHz is a cost-effective RFID-based solution. The reader module comes with an antenna on the chip and can be powered by a 5V power supply. Power on the module and connect the module's transmit pin to the microcontroller's receive pin. Show

ISSN: 2366-1313

your card within the scanning distance, and the card type will be discarded. Optionally, the unit can be configured for additional weight and output.

**PIN of DS3231 RTC module**

As mentioned above, DS3231 IC and 24C32 EEPROM IC are the two main components in a normal DS3231 RTC module board. Other than that, there are a few other components, such as a power LED, some resistors, capacitors, a battery holder, and pins to connect to the microcontroller. Figure 2 shows the components and pins on the RTC DS3231 module.

**SYSTEM FLOW**

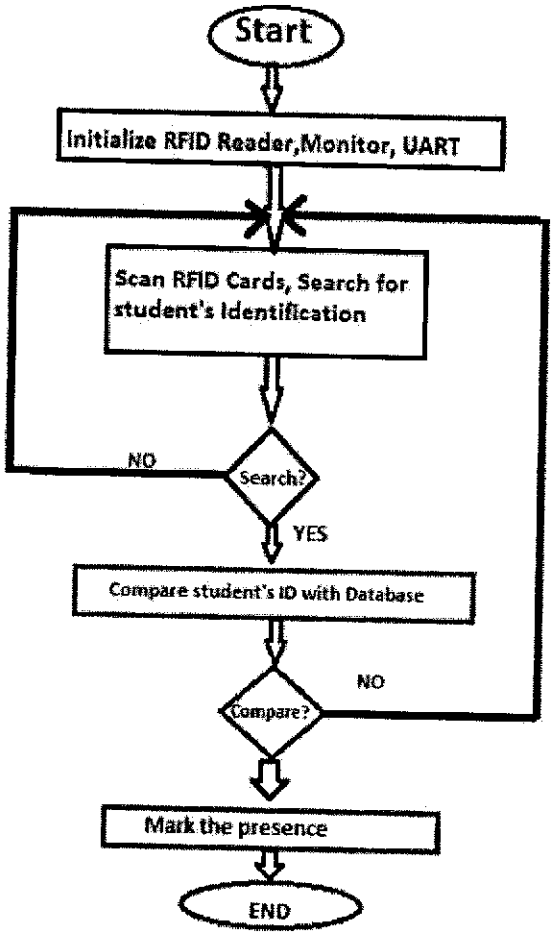



**Fig.1** Flow chart

The proposed attendance monitor uses the idea of IoT to record and obtain information on the server/cloud and make the user obtain it anytime and anywhere. For target assignments, we also like to provide college students access to their attendance, who will log in and view their attendance remotely. We can fully integrate the device with the mobile application so that all functions are from the cell phone itself. Also, we'd like to integrate this device with Canvas or Blackboard using an XML interface.

ISSN: 2366-1313

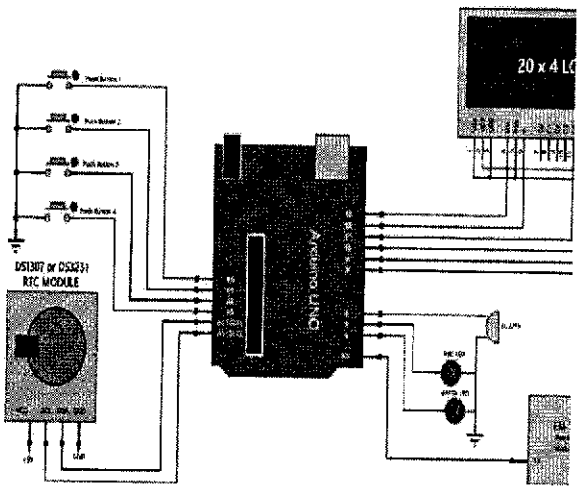



**Fig.2**Circuit connection of the project

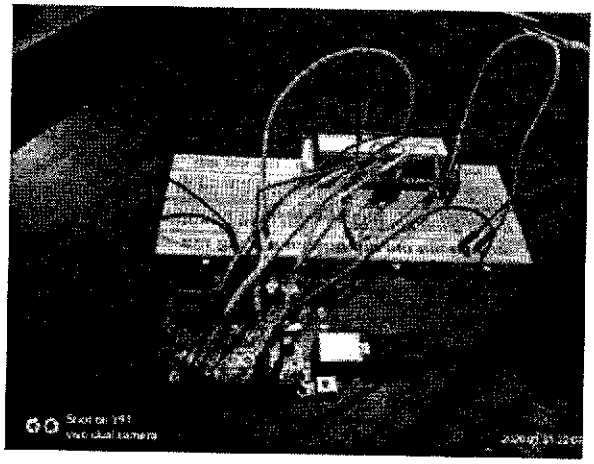


**Fig.3**Experimental results

## IV. RESULTS AND DISCUSSIONS

In this stage, we note some limitations and discuss our system's destination plans. For this prototype, due to laboratory limitations, we invited only five volunteers to participate in our experiments and evaluated the overall performance of the device on this basis. However, detection accuracy may be affected by greatly increasing diversity in humans. This is because more human beings; are likely to have similar framework features, so it may require us to have more granular features. Moreover, real-time capability can also be a key consideration to further improve the robustness of our devices. The fourth figure indicates the circuit's test run and provides our task's real-time output.

## CONCLUSION

Our goal is to increase easy, portable, and facility-friendly RFID-based assistance. The machine features a natural and environment-friendly response to monitoring student attendance on a large scale. The proposed attendance tracking system takes advantage of the Internet of Things to record and retrieve information on the server/cloud and make it available to the consumer anytime and anywhere. For future work, we'd also like to give students access to their attendance, so they can log in and prove their attendance remotely. We can fully integrate the device with the cell phone application so that all capacity is in the cell itself. Also, we'd like to integrate this system with Canvas or Blackboard using an XML interface.

## REFERENCES

ISSN: 2366-1313

1. Sri Madhu BM and Kavya K, 2017, "IoT based automatic attendance management system, IEEE.
2. Qianwen M, Ruchan W, 2018, Smart Attitude Algorithm with Parrite RFID Tags. IEEE.
3. Ertan Z, Ghassan AH, 2018, "New anti-collision protocol for RFID – Based Student Attendance system".
4. Khoirun N, Roudhotul AS et al, 2019, "Implementation of RFID Attendance System With Face Detection Using Validation Viola-Jones and Local Binary Pattern Histogram Method" . (ISESD). October 2019
5. Nguyen HK, Chew MT. "RFID – Based Attendance management System", 2nd Workshop on Recent Trends in Telecommunications Research (RTTR) Palmerston North, New Zealand. 2017.
6. Sri vignesh PSS, Bhaskar M. "RFID and Pose Invariant Face Verification Based Automated Classroom Attendance System" . International Conference on Microelectronics, Computing and Communications (MicroCom). January 2016.
7. Wang Z, Wang R, "One More Tag Enables Fine – Grained RFID Localization and Tracking". IEEE/ACM Transactions on Networking 2019; 26:161-78
8. Sri Madhu, 2017, "IoT based Automatic Attendance Management System", pp.83-86
9. O. Shoewn Development of Attendance Management System using Biometrics. The Pacific Journal of Science and Technology Volume 13, No 1, May 2012
10. Prasadu Peddi (2018), "A STUDY FOR BIG DATA USING DISSEMINATED FUZZY DECISION TREES", ISSN: 2366- 1313, Vol 3, issue 2, pp: 46-57.
11. W.Zhao, R. Chellapa, P.J.Phillips and A.Rosenfld,IFace Recognition: A Literature Survey,vol. 35, No 4, Dec 2003, pp.399-458.

JOURNAL OF HARBIN INSTITUTE OF TECHNOLOGY
ISSN: 0367-6234 Vol. 54 Iss. 9 2022

# KEY GENERATION USING GENETIC ALGORTHM FOR DNA PLAYFAIR CRYPTOSYSTEM

**G. Ramaswamy[*1] & Dr. R. Satya Prasad[2]**
[*1]Research Scholar, Acharya Nagarjuna University
[2]Professor in the Department of Computer Science and Engineering, Acharya Nagarjuna University

**ABSTRACT**

One kind of encryption that makes use of DNA to encode and decode data is called "DNA cryptography." DNA sequences, with their peculiar arrangement of information, may be used to easily apply security measures for sensitive data. Though much has been studied and numerous algorithms have been created for concealing the data, DNA sequence based data encryption seems to be an efficient option for addressing the existing information security demands. The work applies the classic Playfair cipher to DNA cryptography. There are 26 different alphabets, however only 25 are used in the key matrix of a standard matrix since the key matrix is only 5 by 5. In this study, we employ genetic algorithms to produce the key, and an 8x8 key matrix containing 64 DNA codons to increase security. Although the encryption method has been improved, it still follows some of the fundamental principles of the original Playfair Algorithm. To address these issues and strengthen the security of the Playfair method, the DNA codons are rearranged into an 8x8 matrix. The suggested approach is time-effective since it makes use of a LOOKUP table with 64 possible values (A..Z, a..z, 0..9, etc). Additionally, each transmission has a unique LOOKUP table arrangement. We implement the suggested approach and compare its runtime complexity to that of several well-known ciphers. To that end, this paper's primary contribution is a DNA-based security mechanism.

**Keywords:** Deoxyribonucleic acid, Genetic Algorithm, Encryption, Decryption.

**DOI:**10.11720/JHIT.54092022.10

## 1. INTRODUCTION

Information security is a fundamental part of CS. Technology has raised hazards. New transmission algorithms are needed. Humans made an invention that makes transmitted data illegible, so the attacker doesn't know. The attacker attempts several methods to crack the algorithms. Cryptography ensures privacy, authentication, and data integrity. Encryption and decryption transform plaintext to ciphertext. Encryption and decryption need keys. Key type determines security mechanism. Using the same key for encryption and decryption is symmetric key cryptography. AsymmetricKey Cryptography uses a public key for encrypting and a private key for decrypting. As cryptographic algorithms are broken, researchers work to design secure ones. DNA computing involves processing DNA sequences. DNA may be used to hide data in DNA cryptography, according to Adleman (1994) and Gehani et al. (2003). Along with traditional cryptography, new high-security methods are introduced. Information is stored in a DNA sequence, then computed. By synthesizing DNA, he synthesized all conceivable paths faster than the old method. This is termed huge parallelism. Parallelizing DNA helped researchers address NP Hardproblems. Adleman inspired Lipton to use DNA to address the SAT satisfiability issue. DNA computers hold 700 terabytes of data in 1 gram of DNA, as George Church's team at Harvard University verified [5]. Cryptography in biology aims to create unbreakable algorithms. These algorithms lack a theoretical approach, which makes modeling excellent DNA cryptographicschemes difficult [6]. Playfair cipher[9] , a symmetric encryption algorithm, which takes two nearby characters in the plaintext as single units and changes over them into cipher text. In the initial step the key is processed i.e., expel all the rehashed characters from the key and arrange the remaining key characters in a predefined 5x5 letter set framework (Table.1). There are 26 letter sets in English and these letters in order can be organized in 5x5 matrix and in one cell i and j characters are set. At that point process the plaintext i.e., transform the plaintext to its equivalent DNA bases. Divide the bases into three bases each, called as Codon. Consider the codons pairwise. If the pair has the same codons then insert dummy codon in between them to make all the pairs equal. Even after inserting the dummy codon, still in case the pairs are not equal we pad another codon at fag end. Apply the following play fair rules. The current study offers DNA-based security based on the classic play fair cipher. The key generated using genetic algorithm.

**Table 1:** Key matrix of Traditional Playfair algorithm

| | | | | |
|---|---|---|---|---|
| A | B | C | D | E |
| F | G | H | I/J | K |
| L | M | N | O | P |
| Q | R | S | T | U |
| V | W | X | Y | Z |

- If the characters show up on a similar line of the grid then they are replaced by the successive right characters.
- If the characters show up on a similar section of the framework then they are displaced by the immediate underneath characters.
- If the characters are in various rows and columns, a square shape is framed joining the letters of the plaintext. The characters at the corners are supplanted by the other match of corners of the square shape characterized by the actual pair.

In this Playfair method, the key and plaintext limit to 26 characters only and there is no probability of taking any extraordinary character set. So as to overcome this, two authors, Mona Sabry and Atito proposed DNA Playfair which is restricted to 26 characters by considering the codons. Table 2 and 3 shows DNA digital coding and Lookuptable.

Methods used in DNA cryptography are described in Section 2. The creation of keys is covered in Section 3, and the suggested algorithmic schemes are presented in Section 4. Section 5 covers the results of the experiments and analyzes the efficiency of the suggested approach, and Section 6 wraps everything up by discussing the conclusions and the potential extensions of the work.

**Table 2:** DNA Digital Coding

| DNA Base | Binary Form |
|---|---|
| A | 00 |
| C | 01 |
| G | 10 |
| T | 11 |

**Table 3:** Lookup Table

| S.No. | DNA Codon | Replaceable Character | S.No. | DNA Codon | Replaceable Character | S.No. | DNA Codon | Replaceable Character |
|---|---|---|---|---|---|---|---|---|
| 1 | TTT | A | 22 | CCC | V | 43 | AAA | q |
| 2 | TTC | B | 23 | CCA | W | 44 | AAG | r |
| 3 | TTA | C | 24 | CCG | X | 45 | AGT | s |
| 4 | TTG | D | 25 | CAT | Y | 46 | AGC | t |
| 5 | TCT | E | 26 | CAC | Z | 47 | AGA | u |
| 6 | TCC | F | 27 | CAA | a | 48 | AGG | v |
| 7 | TCA | G | 28 | CAG | b | 49 | GTT | w |
| 8 | TCG | H | 29 | CGT | c | 50 | GTC | x |
| 9 | TAT | I | 30 | CGC | d | 51 | GTA | y |
| 10 | TAC | J | 31 | CGA | e | 52 | GTG | Z |
| 11 | TAA | K | 32 | CGG | f | 53 | GCT | 0 |
| 12 | TAG | L | 33 | ATT | g | 54 | GCC | 1 |
| 13 | TGT | M | 34 | ATC | h | 55 | GCA | 2 |
| 14 | TGC | N | 35 | ATA | i | 56 | GCG | 3 |
| 15 | TGA | O | 36 | ATG | j | 57 | GAT | 4 |
| 16 | TGG | P | 37 | ACT | k | 58 | GAC | 5 |
| 17 | CTT | Q | 38 | ACC | l | 59 | GAA | 6 |
| 18 | CTC | R | 39 | ACA | m | 60 | GAG | 7 |
| 19 | CTA | S | 40 | ACG | n | 61 | GGT | 8 |
| 20 | CTG | T | 41 | AAT | o | 62 | GGC | 9 |
| 21 | CCT | U | 42 | AAC | p | 63 | GGA | . |
| | | | | | | 64 | GGG | SPACE |

## 2. RELATED WORKS

Gambir Singh, the authors proposed Cryptography based on DNA which was used the substitution method on using distinct One-Time Pad libraries.They performed molecular computation and indexed using XOR scheme with random key strings. They assumed short segments of unique plaintext messages as input and maintained a codebook to convert the plaintext segment into encrypted data. Using random one-time pad cryptosystem makes the algorithm unbreakable. They also presented DNA Chip using micro-array technology for 2D input and output data. They examinednumerous DNA based steganography systems, which tag the input DNA in secret and then hide it within collections of other DNA. They discussed various modified DNA steganography systems which improved the security[15].

**Hassan et.al** proposed a novel DNA cryptography method created on central dogma on molecular biology. The central dogma is used to produce the DNA into Protein which will be useful for our body. In this process splicing mechanism is used. In this method the author used the splicing mechanism to cut out the fake DNA sequences and join the extrons which is nothing but mRNA. This mRNA sequence converted into protein sequence which then is transmitted to receiver. The receiver performed the reverse process and gets the plaintext. Here, the key is generated randomly and for each communication new key was generated[9].

**Suthar et.al** implemented YAEA encryption algorithm to provide security to the information using DNA cryptography. He downloaded the CanisFamiliaris genome from the GeneBank. It can be transmitted through secure media to the receiver. Transform the plaintext message or images into DNA sequences. Search algorithm is applied to find the position of quadruple DNA sequence in the CanisFamiliaris genome. The same process is done for all the plaintext characters and store the random location pointers into a separate file which is cipher text[18].

**Vinay Cui** designed an encryption scheme by using DNA digital coding, PCR amplification and DNA synthesis and also use the traditional cryptography theory. In this algorithm, the generation of key is the main security aspect and the key pair is PCR primers. The two pairs of primers not designed by the sender or receiver. With the cooperation of both sender and receiver the correct primer sequence is identified. The plaintext is converted into hexadecimal and in turn converted into binary form. This data is encrypted using receiver's public key and then converted into binary cipher data and then into DNA sequences. The sender by mixing the cipher text with certain fake DNA sequences a final sequence is generated which is sent to the receiver through open communicational channel. In the receiver end the receiver takes out the secret message from the mixture of ciphertext and fake data by using correct primer pair. Digital DNA coding is applied and decrypted the binary to plaintext by using his secret key[14].

**Beenish Anam, zhang** discussed various DNA Cryptographic algorithms and also differentiated modern cryptography, quantum cryptography and DNA cryptography. He emphasised all the DNA technologies like PCR, Primers, DNA Chip technology and Steganography using DNA. The author discussed various schemes like the public key cryptography RSA with DNA technology and different operations performed in DNA sequences like annealing, ligation and XOR mechanism[10].

**Harneet Singh** proposed a secure mobile networks approach using DNA Cryptography. In this algorithm, the author converted the plaintext messages into DNA sequences. The author generated two keys from this DNA sequence. For the first Key, find the randomly selected DNA sequence which is called as intron and find the number of times the intron present in the DNA sequence and remove all the identified position intron sequences. The second key value is mapping the codon into amino acid and store the flag value into Key file[13].

**Monica Borda**, Santh et.al proposed Secret Writing Techniques using DNA. In this algorithm the author used XOR and Chromosome DNA indexing. The OTP is generated by using MATLAB bioinformatics toolbox. In this algorithm the author used two different techniques one is steganography and other one is encryption. In the steganography technique, take a plaintext and converted into binary form. Later, generate the OTP by using MATLAB bioinformatics toolbox with the minimum length of binary bits x 10. Each binary value has a corresponding oligonucleotide sequence and is the complementary from the randomly generated sequence. In chromosomes DNA indexing, the author get the sequence from chromosome X which was publicly available database[17].

**Calina Popovici Thangual**, described about the Cryptography, advantages of DNA cryptography. He also implemented an algorithm which uses RSA algorithm for key generation. This algorithm works for both text and images, first convert the given input message into ASCII Code and then convert all the numerical values into a string. Later, the number obtained is encrypted with the public key of the receiver and got another series of numbers. After that, convert these series of numbers into binary and then converted into a DNA sequence, which is the cipher text. The algorithm is strong enough because of using RSA Key generation cryptography algorithm[10].

**Mona Sabry et al** designed play fair cipher implementation using amino acids. In this model, the author takes the plaintext and transforms into binary form. This binary form data turned into DNA sequences. The DNA sequences can be mapped into DNA codons and then turned into Amino acid. In the conversion of codon to amino acid there is an ambiguity and also we have only twenty alphabet values from the 64 codons. The author distributed all the 64 codons to English alphabet set. The author overcame the ambiguity problem by sending the ambiguity number to the receiver[4].

**Qinghai Gao** proposed DNA based encryption system using biological alphabets. In DNA, the genetic information contains only 4 bases that is A, C, G and T. The author also proposed key distribution algorithm which holds one secret code book transmitted via secure channel that has been agreed by both sender and receiver. The sender identified a sequence which match the codebook maximally and send through public channel to the receiver. Convert the plaintext into binary and then convert into DNA sequence, map these DNA sequence into RNA sequence and then Protein Sequence. According to central dogma of molecular biology one way is possible that is DNA to Protein but the reverse is not suitable. But here, the author identified the redundancy and to overcome it he proposed more rounds of Protein to Protein sequence. Reverse translation of protein sequence reproduce two thirds of RNA Sequences. When receiver receives the sequence and find the matches, retrieve unmatched letters. By using the chemical bonding of DNA sequences the author implemented steganography[16].

## 3. KEY GENERATION

**Genetic algorithm**

Genetic algorithms [5] use natural selection to direct randomized search and optimization. Selection, crossover, and mutation are genetic algorithm operators. GA cycles via selection, crossover, and mutation until stopping requirements are met. Reproduction and crossover help genetic algorithms seek.

A. Selection: It's a quantitative method that selects chromosomes from populations depending on their fitness.
B. Crossover: In a crossover surgery, two chromosomes are extracted and a new one is created by combining their traits. After the third locus in each string, 11110111 and 01001010 might be produced. Single-point, two-point, and uniform crossover are crossover procedures.
C. Mutation: Mutation maintains genetic variety between population generations. It's a mutation. GAs modify solution elements string-based. These feature "Gas" bit-reversal. Randomly swap two bits or flip chromosomal bits. Mutating the fifth position of 00000111 yields 00001111.

**Methodology**

Initial chromosome generation is a random hexadecimal number. 128-bit beginning population. Generates 'n' persons. These people attend a fitness event. This fitness function is a maxima function, therefore the fittest person advances. After this, we choose two winners. On chosen people, a random integer determines the crossover point. We acquire the chosen individuals' children after crossing. Again, the fitness function is applied to the offspring, and if they're fitter than the parents, they replace them. The previous step's output becomes the mutation's input. We'll acquire the encryption key after mutation. The essential genetic population generation stages are:

Generation 1:A decimal random number generator generates 128-bit chromosomal populations.
Conversion: Decimal to hexadecimal.

Calculating fitness: Individual fitness is calculated. The repeated maximum symbol is used to calculate fitness. The whole procedure is repeated hundreds of times. Throughout each cycle, the population with the highest fitness score is recorded. Once the termination condition is fulfilled, the population with the highest fitness value is chosen as the encryption key. The essential key Generation Process is shown in Figure 1. The largest population is treated as a key and converted it into decimal number, which is then transformed into a string using the corresponding ASCII values.

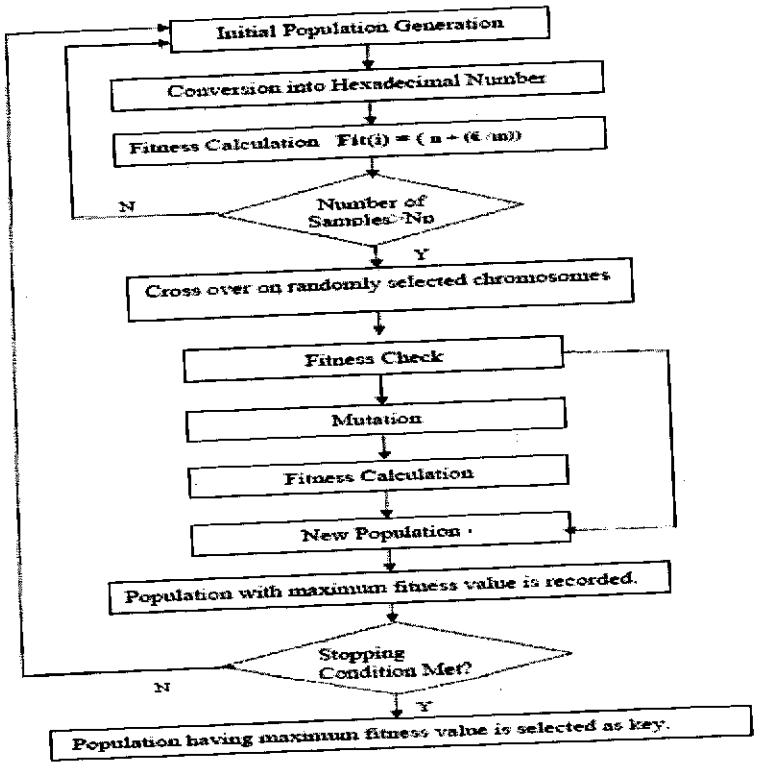


**Figure 1:** Key Generation Process using Genetic algorthm

## 4. PROPOSED ALGORITHM

To eliminate the problem and enhance the security, the proposed model is redesigned by extending the key matrix from 5x5 to 8x8. The key matrix(Table 4.2) originally contained all 64 codons into 8x8 matrix and must be agreed by both sender and receiver. One more strong point in this proposed cryptosystem is the key value that lies between the character set {a..z,A..Z,0..9,space,fullstop}. In traditional playfair cipher the key contains only any 26 characters but in this proposed scheme it extended from 26 to 64. These key values can be mapped into LOOKUP table(Table 3.1). First, we have to process the key value by using following steps.

1. Remove the repeated characters from the key.
2. Map each character with equivalent DNA codon from the LOOKUP table.
3. Arrange the codons that received after processing key into key matrix as first row and rearrange the matrix.

**Table 4:** Key matrix of DNA Playfair Cryptosystem

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| TTT | TAT | ATT | CAT | CTT | AAT | GTT | GAT |
| TTC | TAC | ATC | CAC | CTC | AAC | GTC | GAC |
| TCT | TGT | ACT | CGT | CCT | AGT | GCT | GGT |
| TTG | TAG | ATG | CAG | CTG | AAG | GTG | GAG |
| TTA | TAA | ATA | CAA | CTA | AAA | GTA | GAA |
| TCC | TGC | ACC | CGC | CCC | AGC | GCC | GGC |
| TCA | TGA | ACA | CGA | CCA | AGA | GCA | GGA |
| TCG | TGG | ACG | CGG | CCG | AGG | GCG | GGG |

Later, transform the plaintext into its equivalent DNA bases. Divide these bases into codons and consider the codon pair for applying traditional playfair.

**Algorithm for Encryption**

***Algorithm Encryption(Message,K)***
Message is the Plaintext and K is the Key value

**Begin**

1. Arrange 64 Codons into 8x8 matrix
2. Map each character of the key to the LOOKUP table(Table 3.1) and place the codons in the key matrix and rearrange the key matrix.
3. Transform the Message into DNA Sequences.
4. Divide DNA Sequences into Codons.
5. If there exists two codons consequently insert dummy codon in between them
6. If the Pairs of codons are not even pad another dummy codon to make it as even.
7. After apply playfair rules for each codon pair and get Cipher text

**End**

**Algorithm for Decryption**

***Algorithm Decryption(C)***
C is the Cipher text. K is the key value and key matrix received from the sender through secure medium
**Begin**

1. Map each character of the key to the LOOKUP table(Table 3) and place the codons in the key matrix and rearrange the key matrix.
2. Divide the sequence into Codons and apply playfair rules.
3. Remove the padded Codons, if necessary
4. Transform the DNA sequences into equivalent ASCII Character which is plaintext.

**End**

**Empirical Analysis**

***Process of Encryption***
Let plaintext M=communication and Key k=suspect. In the first step, process the key which means remove the repeated characters from it. In the present example s is repeated twice. After removing s Key value k=supect.In the next step, map each character with the LOOKUP table(Table 3).

s-AGT, u-AGA, p-AAC, e-CGA, c-CGT, and t-AGC
Now, arrange these values in to key matrix and rearrange the key matrix.

**Table 5:** Key matrix after formatting

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| AGT | AGA | AAC | CGA | CGT | AGC | TTT | TAT |
| ATT | CAT | CTT | AAT | GTT | GAT | TTC | TAC |
| ATC | CAC | CTC | GTC | GAC | TCT | TGT | ACT |
| CCT | GCT | GGT | TTG | TAG | ATG | CAG | CTG |
| AAG | GTG | GAG | TTA | TAA | ATA | CAA | CTA |
| AAA | GTA | GAA | TCC | TGC | ACC | CGC | CCC |
| GCC | GGC | TCA | TGA | ACA | CCA | GCA | GGA |
| TCG | TGG | ACG | CGG | CCG | AGG | GCG | GGG |

After key processing, process the plaintext. In the process of plaintext, first convert the characters into their equivalent DNA Sequences.

**Table 6:** Secret Message Binary Coding

| Character | ASCII VALUE | Binary Value | DNA Digital Coding |
|---|---|---|---|
| f | 102 | 01100110 | CGCG |
| i | 105 | 01101001 | CGGC |
| r | 114 | 01110010 | CTAG |
| s | 115 | 01110011 | CTAT |
| t | 116 | 01110100 | CTCA |
| m | 109 | 01101101 | CGTC |
| e | 101 | 01100101 | CGCC |
| s | 115 | 01110011 | CTAT |
| s | 115 | 01110011 | CTAT |
| a | 97 | 01100001 | CGAC |
| g | 103 | 01100111 | CGCT |
| e | 101 | 01100101 | CGCC |

The DNA sequence of the plaintext is,
CGCGCGGCCTAGCTATCTCACGTCCGCCCTATCTATCGACCGCTCGCC
Divide the sequence into codons and take pairwise data to apply the playfair rule then
CGC,GCG|GCC,TAG|CTA,TCT|CAC,GTC|CGC,CCT|ATC,TAT|CGA,CCG|CTC,GCC

If the number of DNA sequences count is divisible by three then no need to pad the DNA sequence. If the number of DNA sequences count is divisible by three and leaves a remainder 1 then pad AA to the sequence to make the pair. If the number of DNA sequences count is divisible by three and leaves a remainder 2 then pad A to the sequence to make the pair.In this example the total count of sequences divisible by 3 and leaves a remainder 1 so pad AA to the sequence then,
CGC,GCG|GCC,TAG|CTA,TCT|CAC,GTC|CGC,CCT|ATC,TAT|CGA,CCG|CTC,GCC
Consider the first pair CGC, GCG and find the positions in the Keymatrix and apply the playfair rule.

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| AGT | AGA | AAC | CGA | CGT | AGC | TTT | TAT |
| ATT | CAT | CTT | AAT | GTT | GAT | TTC | TAC |
| ATC | CAC | CTC | GTC | GAC | TCT | TGT | ACT |
| CCT | GCT | GGT | TTG | TAG | ATG | CAG | CTG |
| AAG | GTG | GAG | TTA | TAA | ATA | CAA | CTA |
| AAA | GTA | GAA | TCC | TGC | ACC | CGC | CCC |
| GCC | GGC | TCA | TGA | ACA | CCA | GCA | GGA |
| TCG | TGG | ACG | CGG | CCG | AGG | GCG | GGG |

The pair is in same column value, replace with successive down values . Then the values are GCA, TTT. Proceeding the same for the remaining pairs.
For the pair GCC,TAG the result is ACA,CCT

| AGT | AGA | AAC | CGA | CGT | AGC | TTT | TAT |
|---|---|---|---|---|---|---|---|
| ATT | CAT | CTT | AAT | GTT | GAT | TTC | TAC |
| ATC | CAC | CTC | GTC | GAC | TCT | TGT | ACT |
| CCT | GCT | GGT | TTG | TAG | ATG | CAG | CTG |
| AAG | GTG | GAG | TTA | TAA | ATA | CAA | CTA |
| AAA | GTA | GAA | TCC | TGC | ACC | CGC | CCC |
| GCC | GGC | TCA | TGA | ACA | CCA | GCA | GGA |
| TCG | TGG | ACG | CGG | CCG | AGG | GCG | GGG |

The next pair CTA,TCT the result is ATA,ACT

| AGT | AGA | AAC | CGA | CGT | AGC | TTT | TAT |
|---|---|---|---|---|---|---|---|
| ATT | CAT | CTT | AAT | GTT | GAT | TTC | TAC |
| ATC | CAC | CTC | GTC | GAC | TCT | TGT | ACT |
| CCT | GCT | GGT | TTG | TAG | ATG | CAG | CTG |
| AAG | GTG | GAG | TTA | TAA | ATA | CAA | CTA |
| AAA | GTA | GAA | TCC | TGC | ACC | CGC | CCC |
| GCC | GGC | TCA | TGA | ACA | CCA | GCA | GGA |
| TCG | TGG | ACG | CGG | CCG | AGG | GCG | GGG |

The next pair, CAC,GTC , these are also in the same row, take the immediate right codon values the result is CTC,GAC

| AGT | AGA | AAC | CGA | CGT | AGC | TTT | TAT |
|---|---|---|---|---|---|---|---|
| ATT | CAT | CTT | AAT | GTT | GAT | TTC | TAC |
| ATC | CAC | CTC | GTC | GAC | TCT | TGT | ACT |
| CCT | GCT | GGT | TTG | TAG | ATG | CAG | CTG |
| AAG | GTG | GAG | TTA | TAA | ATA | CAA | CTA |
| AAA | GTA | GAA | TCC | TGC | ACC | CGC | CCC |
| GCC | GGC | TCA | TGA | ACA | CCA | GCA | GGA |
| TCG | TGG | ACG | CGG | CCG | AGG | GCG | GGG |

The next pair CGC, CCT. The result is AAA, CAG

| AGT | AGA | AAC | CGA | CGT | AGC | TTT | TAT |
|---|---|---|---|---|---|---|---|
| ATT | CAT | CTT | AAT | GTT | GAT | TTC | TAC |
| ATC | CAC | CTC | GTC | GAC | TCT | TGT | ACT |
| CCT | GCT | GGT | TTG | TAG | ATG | CAG | CTG |
| AAG | GTG | GAG | TTA | TAA | ATA | CAA | CTA |
| AAA | GTA | GAA | TCC | TGC | ACC | CGC | CCC |
| GCC | GGC | TCA | TGA | ACA | CCA | GCA | GGA |
| TCG | TGG | ACG | CGG | CCG | AGG | GCG | GGG |

The next pair ATC,TAT the result is ACT,AGT

| AGT | AGA | AAC | CGA | CGT | AGC | TTT | TAT |
|---|---|---|---|---|---|---|---|
| ATT | CAT | CTT | AAT | GTT | GAT | TTC | TAC |
| ATC | CAC | CTC | GTC | GAC | TCT | TGT | ACT |
| CCT | GCT | GGT | TTG | TAG | ATG | CAG | CTG |
| AAG | GTG | GAG | TTA | TAA | ATA | CAA | CTA |
| AAA | GTA | GAA | TCC | TGC | ACC | CGC | CCC |
| GCC | GGC | TCA | TGA | ACA | CCA | GCA | GGA |
| TCG | TGG | ACG | CGG | CCG | AGG | GCG | GGG |

The next pair CGA,CCG. the result would be CGG,CCT

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| AGT | AGA | AAC | CGA | CGT | AGC | TTT | TAT |
| ATT | CAT | CTT | AAT | GTT | GAT | TTC | TAC |
| ATC | CAC | CTC | GTC | GAC | TCT | TGT | ACT |
| CCT | GCT | GGT | TTG | TAG | ATG | CAG | CTG |
| AAG | GTG | GAG | TTA | TAA | ATA | CAA | CTA |
| AAA | GTA | GAA | TCC | TGC | ACC | CGC | CCC |
| GCC | GGC | TCA | TGA | ACA | CCA | GCA | GGA |
| TCG | TGG | ACG | CGG | CCG | AGG | GCG | GGG |

The next pair CTC,GCC then the result is TCA,ATC

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| AGT | AGA | AAC | CGA | CGT | AGC | TTT | TAT |
| ATT | CAT | CTT | AAT | GTT | GAT | TTC | TAC |
| ATC | CAC | CTC | GTC | GAC | TCT | TGT | ACT |
| CCT | GCT | GGT | TTG | TAG | ATG | CAG | CTG |
| AAG | GTG | GAG | TTA | TAA | ATA | CAA | CTA |
| AAA | GTA | GAA | TCC | TGC | ACC | CGC | CCC |
| GCC | GGC | TCA | TGA | ACA | CCA | GCA | GGA |
| TCG | TGG | ACG | CGG | CCG | AGG | GCG | GGG |

The final cipher text is

**GCATTTACACCTATAACTCTCGACAAACAGACTAGTCGGCCTTCAATC**

***Process of Decryption***

First, the receiver process the value of key received from a secured channel. And divide the cipher text into pairwise codons.

**GCA,TTT|ACA,CCT|ATA,ACT|CTC,GAC|AAA,CAG|ACT,AGT|CGG,CCT|TCA,ATC|**

The receiver also process the key by following the procedure as sender done then the key matrix is,

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| AGT | AGA | AAC | CGA | CGT | AGC | TTT | TAT |
| ATT | CAT | CTT | AAT | GTT | GAT | TTC | TAC |
| ATC | CAC | CTC | GTC | GAC | TCT | TGT | ACT |
| CCT | GCT | GGT | TTG | TAG | ATG | CAG | CTG |
| AAG | GTG | GAG | TTA | TAA | ATA | CAA | CTA |
| AAA | GTA | GAA | TCC | TGC | ACC | CGC | CCC |
| GCC | GGC | TCA | TGA | ACA | CCA | GCA | GGA |
| TCG | TGG | ACG | CGG | CCG | AGG | GCG | GGG |

Consider the first pair GCA, TTT and find the positions in the Keymatrix and apply the play fair rule.The result is CGC,GCG

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| AGT | AGA | AAC | CGA | CGT | AGC | TTT | TAT |
| ATT | CAT | CTT | AAT | GTT | GAT | TTC | TAC |
| ATC | CAC | CTC | GTC | GAC | TCT | TGT | ACT |
| CCT | GCT | GGT | TTG | TAG | ATG | CAG | CTG |
| AAG | GTG | GAG | TTA | TAA | ATA | CAA | CTA |
| AAA | GTA | GAA | TCC | TGC | ACC | CGC | CCC |
| GCC | GGC | TCA | TGA | ACA | CCA | GCA | GGA |
| TCG | TGG | ACG | CGG | CCG | AGG | GCG | GGG |

For the pair ACA,CCT the result is GCC, TAG

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| AGT | AGA | AAC | CGA | CGT | AGC | TTT | TAT |
| ATT | CAT | CTT | AAT | GTT | GAT | TTC | TAC |
| ATC | CAC | CTC | GTC | GAC | TCT | TGT | ACT |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| **CCT** | GCT | GGT | TTG | TAG | ATG | CAG | CTG |
| **AAG** | GTG | GAG | TTA | TAA | ATA | CAA | CTA |
| **AAA** | GTA | GAA | TCC | TGC | ACC | CGC | CCC |
| **GCC** | GGC | TCA | TGA | ACA | CCA | GCA | GGA |
| **TCG** | TGG | ACG | CGG | CCG | AGG | GCG | GGG |

The next pair ATA, ACT the result is CTA,TCT

| **AGT** | **AGA** | **AAC** | **CGA** | **CGT** | **AGC** | **TTT** | **TAT** |
|---|---|---|---|---|---|---|---|
| **ATT** | CAT | CTT | AAT | GTT | GAT | TTC | TAC |
| **ATC** | CAC | CTC | GTC | GAC | TCT | TGT | ACT |
| **CCT** | GCT | GGT | TTG | TAG | ATG | CAG | CTG |
| **AAG** | GTG | GAG | TTA | TAA | ATA | CAA | CTA |
| **AAA** | GTA | GAA | TCC | TGC | ACC | CGC | CCC |
| **GCC** | GGC | TCA | TGA | ACA | CCA | GCA | GGA |
| **TCG** | TGG | ACG | CGG | CCG | AGG | GCG | GGG |

The next pair CTC,GAC the result is CAC,GTC

| **AGT** | **AGA** | **AAC** | **CGA** | **CGT** | **AGC** | **TTT** | **TAT** |
|---|---|---|---|---|---|---|---|
| **ATT** | CAT | CTT | AAT | GTT | GAT | TTC | TAC |
| **ATC** | CAC | CTC | GTC | GAC | TCT | TGT | ACT |
| **CCT** | GCT | GGT | TTG | TAG | ATG | CAG | CTG |
| **AAG** | GTG | GAG | TTA | TAA | ATA | CAA | CTA |
| **AAA** | GTA | GAA | TCC | TGC | ACC | CGC | CCC |
| **GCC** | GGC | TCA | TGA | ACA | CCA | GCA | GGA |
| **TCG** | TGG | ACG | CGG | CCG | AGG | GCG | GGG |

The next pair AAA,CAG. The result is CGC,CCT

| **AGT** | **AGA** | **AAC** | **CGA** | **CGT** | **AGC** | **TTT** | **TAT** |
|---|---|---|---|---|---|---|---|
| **ATT** | CAT | CTT | AAT | GTT | GAT | TTC | TAC |
| **ATC** | CAC | CTC | GTC | GAC | TCT | TGT | ACT |
| **CCT** | GCT | GGT | TTG | TAG | ATG | CAG | CTG |
| **AAG** | GTG | GAG | TTA | TAA | ATA | CAA | CTA |
| **AAA** | GTA | GAA | TCC | TGC | ACC | CGC | CCC |
| **GCC** | GGC | TCA | TGA | ACA | CCA | GCA | GGA |
| **TCG** | TGG | ACG | CGG | CCG | AGG | GCG | GGG |

The next pair ACT,AGT. then the result is ATC, TAT

| **AGT** | **AGA** | **AAC** | **CGA** | **CGT** | **AGC** | **TTT** | **TAT** |
|---|---|---|---|---|---|---|---|
| **ATT** | CAT | CTT | AAT | GTT | GAT | TTC | TAC |
| **ATC** | CAC | CTC | GTC | GAC | TCT | TGT | ACT |
| **CCT** | GCT | GGT | TTG | TAG | ATG | CAG | CTG |
| **AAG** | GTG | GAG | TTA | TAA | ATA | CAA | CTA |
| **AAA** | GTA | GAA | TCC | TGC | ACC | CGC | CCC |
| **GCC** | GGC | TCA | TGA | ACA | CCA | GCA | GGA |
| **TCG** | TGG | ACG | CGG | CCG | AGG | GCG | GGG |

The DNA sequence is
CGC,GCG|GCC,TAG|CTA,TCT|CAC,GTC|CGC,CCT|ATC,TAT|CGA,CCG|CTC,GCC
Convert the above sequence into binary

01100110011010010111001001110011011101000110110101100101011100110111001101100111011001 01
Transform the binary values into ASCII in turn into its equivalent character.
01100110, 01101001, 01110010, 01110011, 01110100, 01101101, 01100101, 01110011, 01110011, 01100111, 01100101

M= firstmessage

## 5. EXPERIMENTAL RESULTS

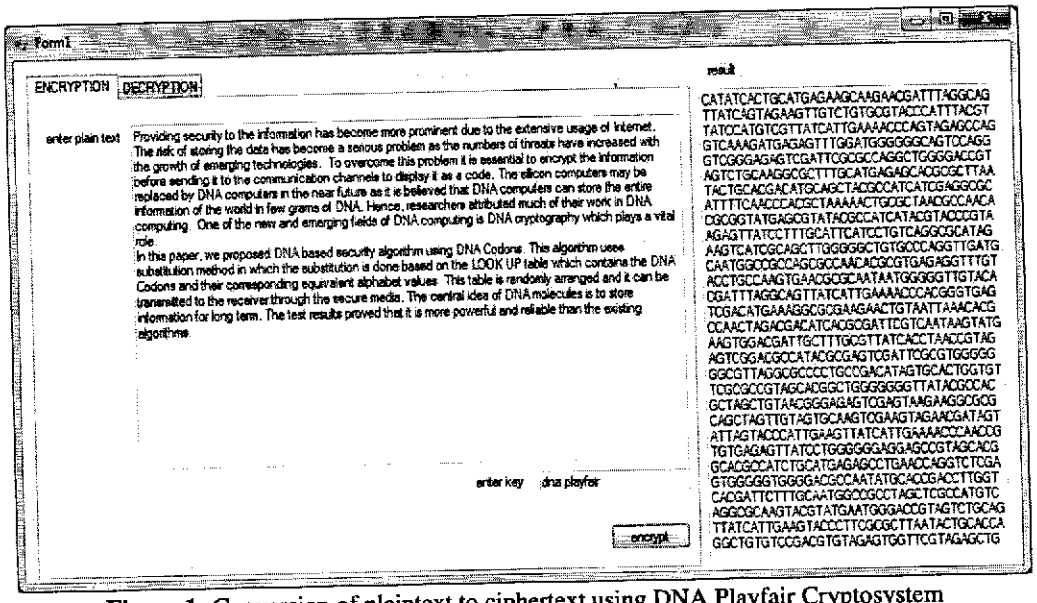



**Figure 1:** Conversion of plaintext to ciphertext using DNA Playfair Cryptosystem

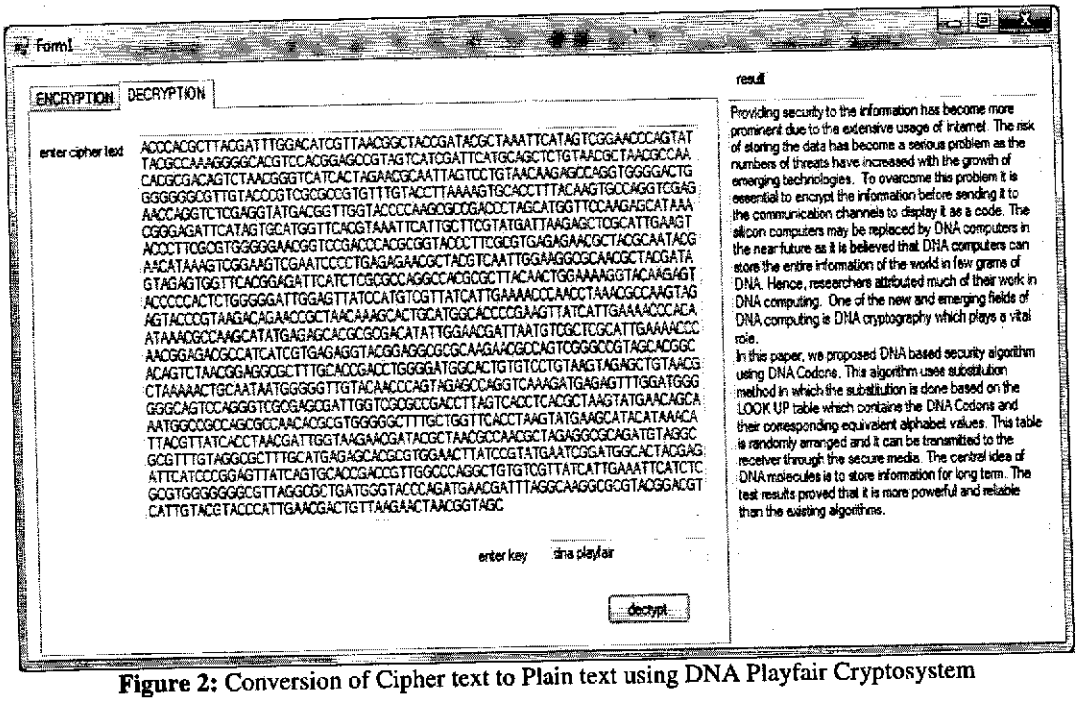



**Figure 2:** Conversion of Cipher text to Plain text using DNA Playfair Cryptosystem

**Performance measurement**

The experimental results were verified in .NET Environment. The performance of proposed algorithm with varying lengths of plaintext were observed for various message lengths of 500,1000,1500,2000 and so on and the key is "dnaplayfair". The length of the message, time for encryption and time for decryption have been tabulated in table 7 and in table 8respectively and depicted in the form of a graph in the figure 3and figure4. The proposed algorithm has taken less time when compared to the playfair algorithm implemented by Atito. The method proposed by Atito has taken more time because of it is not performed the traditional playfair it also stores the ambiguity number .

**Table 7:** Comparison of Encryption Process between Proposed Method and Atito Method for various Message lengths

| S. No. | Number of Characters | Encryption time(in ms) of Proposed Algorithm | Encryption time Proposed by Atito |
|---|---|---|---|
| 1 | 500 | 0.058032 | 0.865411 |
| 2 | 1000 | 0.083312 | 1.870414 |
| 3 | 1500 | 0.159926 | 3.542186 |
| 4 | 2000 | 0.271735 | 4.9856 |
| 5 | 5000 | 1.47053 | 11.62155 |
| 6 | 10000 | 7.478196 | 26.25961 |

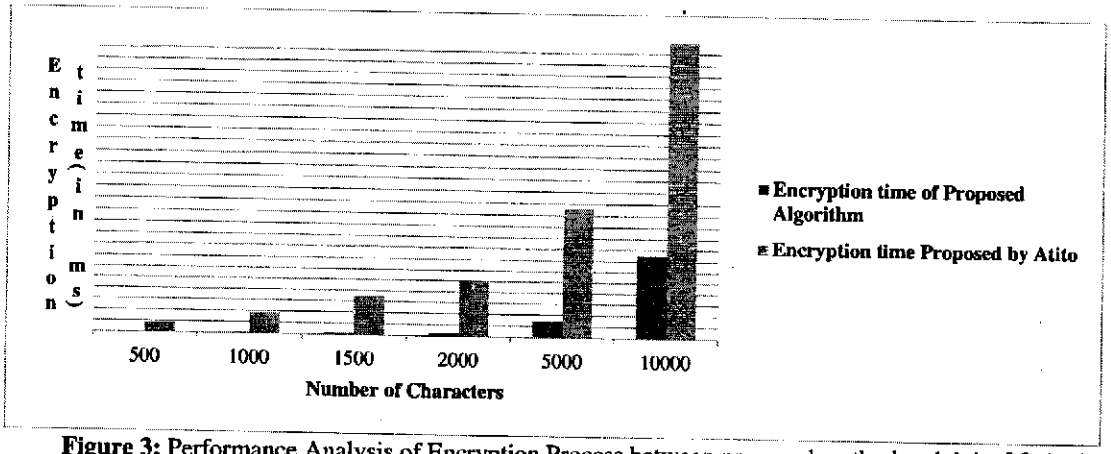



**Figure 3:** Performance Analysis of Encryption Process between proposed method and Atito Method

**Table 8:** Comparison of Decryption Process between Proposed Method and Atito Method for various Message lengths

| Sl.No. | Number of Characters | Decryption time(in ms) of Proposed Algorithm | Decryption time Proposed by Atito |
|---|---|---|---|
| 1 | 500 | 0.0128772 | 0.7659356 |
| 2 | 1000 | 0.0142942 | 1.2345896 |
| 3 | 1500 | 0.0194885 | 3.1254689 |
| 4 | 2000 | 0.0261794 | 4.0356988 |
| 5 | 5000 | 0.0750149 | 10.9841532 |
| 6 | 10000 | 0.1651543 | 19.9856453 |

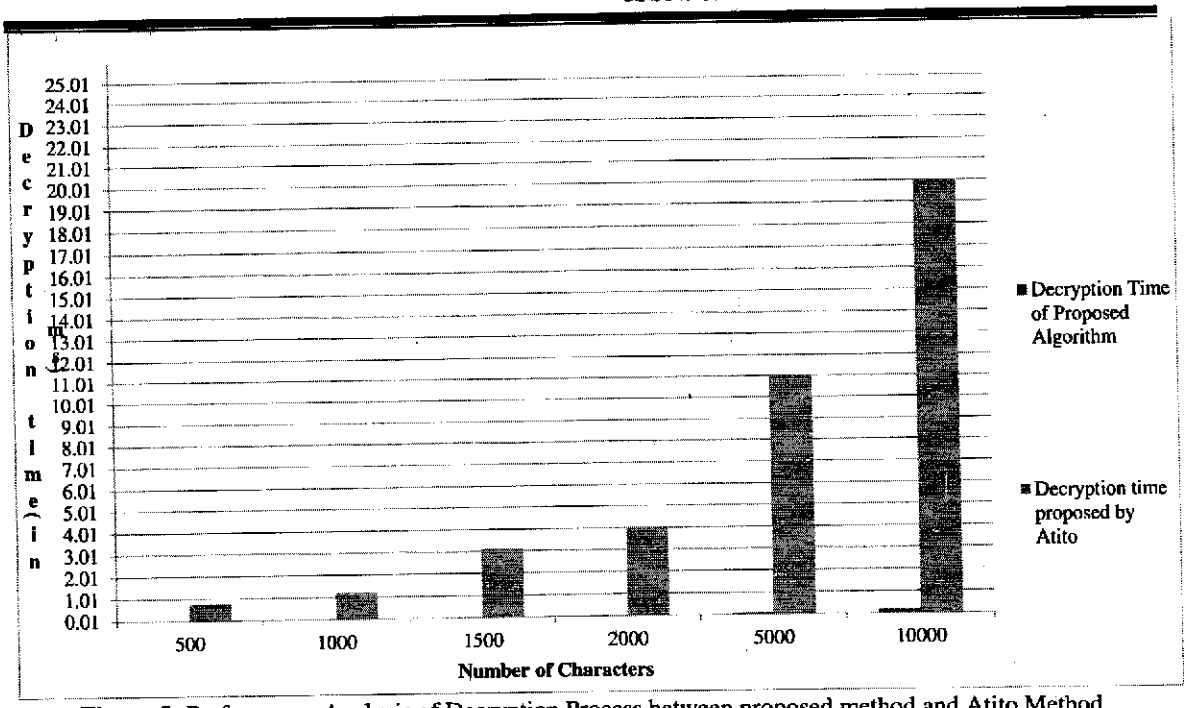



**Figure 5:** Performance Analysis of Decryption Process between proposed method and Atito Method

## 6. CONCLUSION

A day may come when DNA may be consolidated into PC chip to make bio-chip which do not damage to the human life. Not just this, DNA has the high capacity to store tremendous amount of data like terabytes in a single gram of DNA base. The proposed calculation encoded the key dependent on the LOOKUP table and the key not just contains letters, it contains capitalized, small case letters in order, digits and exceptional symbols like space and full stop. Despite the fact that the normal play fair or DNA play fair depends on the 5x5 matrix and the key likewise cut off points to just letters in order characters, the proposed calculation has taken the span of key matrix as 8x8. In view of the outcome investigation, it demonstrates to give greater security. In future it can be implemented using different algorithm for images with no data loss and with the help of new key generation algorithms.

## REFERENCES

[1] Prasanna Balaji Narasingapuram, M. Ponnavaikko, "DNA Cryptography Based User Level Security for Cloud Computing and Applications," International Journal of Recent Technology and Engineering (IJRTE), ISSN: 2277-3878, Volume-8 Issue-5, January 2020.

[2] Hamza Hammami, Hanen Brahmi, Sadok Ben Yahia,"Secured Outsourcing Towards a Cloud Computing Environment Based on DNA Cryptography," IEEE, pp. 31-36, 2018.

[3] Bahubali Akiwate, Latha Parthiban, "Enhanced DNA Cryptographic Solution for Secured Data Transmission," International Journal of Innovative Technology and Exploring Engineering (IJITEE) ISSN: 2278-3075,2019.

[4] Bahubali Akiwate, Latha Parthiban, "A Dynamic DNA for Key-based Cryptography," CTEMS, IEEE, 2018.

[5] Kaur, Karandeep, "A Double Layer Encryption Algorithm based on DNA and RSA for Security on Cloud,"International Research Journal of Engineering and Technology, Volume: 03 Issue: 03 Mar-2016.

[6] Himanshu Kumar Shukla, Satyam Dubey, "Open Access ArticleSecurity in Internet of Things (IoT) Hashing Cryptographic Functions", Vol.7, Issue.3, pp.7-11, Jun-2019.

[7] S.V.Keerthana Priya, S.J.Saritha, "A Robust Technique to Generate Unique Code DNA Sequence," IEEE, pp. 3815-3820, 2017.

[8] Zhang et al, "DNA based random key generation and management for OTP encryption," 2017 Sep;159:51-63. DOI: 10.1016/j.biosystems.2017.07.002. Epub 2017 Jul 18.

[9] Hassan Al-Mahdi Meshrif Alruily Osama R.Shahin, Khalid Alkhaldi, "Design and Analysis of DNA Encryption and Decryption Technique based on Asymmetric Cryptography System," (IJACSA) International Journal of Advanced Computer Science and Applications, Vol. 10, No. 2, 2019.

[10] M. Thangavel, P. Varalakshmi, "Enhanced DNA and ElGamal cryptosystem for secure data storage and retrieval in the cloud," Springer Science+Business Media, LLC, part of Springer Nature 2017. [11] Md.

Rafiul Biswas, Kazi Md. Rokibul Alam, Ali Akber, and Yasuhiko Morimoto, "A DNA Cryptographic Technique Based on Dynamic DNA Encoding and Asymmetric Cryptosystem," IEEE, 978-1-5386-3288-8/17 ©2017.

[11] A.Vyasa Bharadwaja, V. Ganesan, "DNA Computing Based Encryption Algorithm for Wireless Multimedia Communication System," International Journal of Innovative Technology and Exploring Engineering (IJITEE), ISSN: 2278-3075, Volume-9 Issue3, January 2020
Saifali Mavanai, Ajay Pal, Ravi Pandey, Asst Prof. Deepika Nadar, "Message Transmission Using DNA Crypto-System," International Journal of Computer Science and Mobile Computing.

[12] Vinay S, Adarsh Pujar, Ankith, H.Akshay Kedlaya, Vasudev S Shahapur, " Implementation of DNA Cryptography based on Dynamic DNA Sequence Table using Cloud Computing", International Journal of Engineering Research & Technology (IJERT)ISSN: 2278-0181 Published by, RTESIT - 2019 Conference Proceedings

[13] Gambhir Singh, Rakesh Kumar Yadav, "DNA Based Cryptography Techniques with Applications and Limitations,"International Journal of Engineering and Advanced Technology (IJEAT), ISSN: 2249 – 8958, Volume-8 Issue-6, August 2019.

[14] Partha Sarathi Goswami, Tamal Chakraborty, Harekrishna Chatterjee, "A Novel Encryption Technique Using DNA Encoding and Single Qubit Rotations", Vol.6 , Issue.3 , pp.364-369, Mar-2018.

[15] Santhi G., "Secure Data transmission using Recombinant DNA Cryptography and Morse Code Pattern", Vol.8 , Issue.5 , pp.174-177, May-2020.

[16] Jignesh Patel, Foram Suthar, Samrat.V.O.Khanna, "Open Access ArticleA Critical Analysis on Encryption Techniques used for Data Security in Cloud Computing and IOT (Internet of Things) based Smart cloud storage System: A Survey", Vol.7, Issue.2, pp.21-25, Apr-2019.

SEYBOLD Report
ISSN: 1533 - 9211
DOI: 10.5281/zenodo.7135160

# INSPIRED FEISTEL DNA BASED CRYPTOSYSTEM USING D-BOX AND IMAGE BASED KEY GENERATION

**G. Ramaswamy*1 & Dr. R. Satya Prasad*2**
*1Research Scholar, Acharya Nagarjuna University
*2 Professor in the Department of Computer Science and Engineering, Acharya Nagarjuna University

**Abstract:**
DNA cryptography is an exciting development in the growing area of DNA computing. Computing at the molecular level, such as that found in DNA, is known as bio-molecular computing. DNA is not just a data storage medium, but also a computational instrument. Researchers have placed a premium on DNA Computation because to its unique ability to store vast quantities of information. Traditional crypto algorithms have been shown to be flawed. Better information security may be achieved by integrating the fields of biology, chemistry, and computer science into a single, cohesive field of study. In this paper, we present a model that take image as a key and encrypts data by making use of DNA codons. Using a different key created by a key generation method in each cycle is a defining feature of the proposed technique, making it almost difficult to decrypt the picture using a cryptanalytic assault. It is also shown that the suggested method outperforms several preexisting algorithms when assessed by an avalanche effect.
***Keywords: DNA Cryptography, Encryption, Decryption, DNA Codons***

## 1. INTRODUCTION:

Innovations in technology improve levels of safety. In order to earn back customers' faith, several industries, such as banking and emailing, need to improve their security measures. The use of cryptography ensures the data's safety. Using cryptography, one may encrypt text. Encryption is the process of transforming data consisting of text into a message that has not been jumbled, whereas decryption performs the reverse function. The term "ciphertext" refers to a communication that has been decrypted. Quantum cryptography employs photons and can communicate at distances of up to 90 miles [1], whereas modern encryption has trouble with prime factorization]. After that, Raj and his colleagues [2] found the structure that was used to safeguard the data. Both DNA-based substitution and One-Time Pads were developed by him. Using biological components to encrypt data is what DNA cryptography does. It belongs to the realm of computers based on DNA. DNA computing involves the use of DNA sequences rather than traditional silicon processors. To solve NP-complete problems such as the Hamiltonian approach. He illustrated the problem with seven towns named after DNA. By synthesizing DNA, he was able to synthesis all possible pathways more quickly than by using the traditional approach. The name for this phenomenon is "huge parallelism." DNA parallelism was essential in researchers solving NP Hard problems. Using DNA as an example, Lipton instructed Kolte on how to solve SAT questions using the Boolean technique. The goal of cryptography in the

DOI: 10.5281/zenodo.7135160

field of biology is to develop algorithms that cannot be broken[4]. Modeling DNA cryptography systems is made harder as a result of the absence of a theoretical approach in these methods [6]. The present research provides DNA security that is based on the traditional encryption, which makes use of the D-box. The method of encryption changes depending on the required quantity of iterations. The key generator mechanism is used during each cycle to manufacture keys. The process is going to be covered in the next section. In Section 2, we will discuss the many aspects of DNA cryptography. Both Section 3 and Section 4 will discuss the essential generating methods. The experimental data, the performance analysis of the suggested approach, and avalanche effect comparisons with existing methodologies are discussed in Sections 5, respectively. The sixth section serves as the article's conclusion.

## 2. RELATED WORK:

Network security constantly seeks unbreakable cryptographic techniques to safeguard data in the cloud, on a network, or in other situations. Recent years have seen several initiatives. Raj et al. developed two approaches for DNA-based cryptography. One uses OTP, which cannot be broken, while the other uses steganography [2]. Sherif et al. encrypt each plaintext character with four molecules of DNA [11]. Hazra [13] presented a DNA-and carbon nanotube-based cryptosystem. Ponnavaiklu introduced a DNA encryption layer to IDEA. This approach increases the key size to prevent cryptanalysis [14]. Kaushik introduced DNA-Public KeyCryptography, which combines encryption with digital signature [15]. He developed an RSA key-generation algorithm. It works on text and photos. It turns the input message into ASCII code and then into a string. Then, the number is encrypted using the receiver's public key to get another series. Convert these digits to binary, then DNA to get the encrypted text. It employs RSA key generation cryptography .

Bony[6] created an amino acid playfair cipher. This model converts plaintext to binary. Binary data becomes DNA sequences. DNA sequences may be mapped into codons and, subsequently, amino acids. Codon to amino acid conversion is ambiguous as the 64 codons have only 20 alphabetical values. The author added all 64 English codons. The author avoided ambiguity by supplying the ambiguity number .

Tornea presented a DNA-based encryption method. DNA comprises just A, C, G, and T as nucleotides. The author also presented a key distribution mechanism that uses a secure channel and a secret code book. The sender sends a codebook-matching sequence to the recipient via the public channel. Convert plaintext to binary, then DNA. map DNA to RNA and protein. DNA to protein is a molecular biology dogma, but not the opposite. Here, the author detected redundancy and advocated additional protein-to-protein sequence rounds. Protein reverse translation reproduces two-thirds of RNA. The receiver will recover the mismatched letters after receiving the sequence and finding matches. The author constructed steganography utilizing DNA chemical bonding . One researcher encrypts using transposition. They employed fixed-size blocks, and the key matrix size should be the same. They transform plain text into ascii values. Simultaneously, a random key and DNA sequence are created. The transposed

SEYBOLD Report
ISSN: 1533 - 9211
DOI: 10.5281/zenodo.7135160

block value and DNA sequence value form a new matrix. The matrix is then rotated row-by-column. Ascii to characters (ciphertext) [10]. Another researcher suggested encrypting the data using a pixel-based algorithm. First, picture pixels were jumbled. Then they utilize camouflaged graphics and watermarking to obscure the data [8].Nath[11] also developed an image encryption scheme based on RGB pixel displacement. This method extracts the picture's RGB values. Transform the vector to a matrix with the same RGB dimensions as the original picture.

## 3. KEY GENERATION:

G generates a private key from an image seed. Network D must differentiate between the generator's private key and transformation data. Source and transformation domains create keys. The source domain includes "seed" photographs. Transformation domain contains image's destination. The transformation domain reflects the private key's "style," such as a chaotic high-security key. Here's how DeepKeyGen loses keys.

$L = L_G + L_D$ (1)

In the equation, $L_G$ and $L_D$ denote the generator and discriminator loss functions.

1) G-Network alters the picture. Generates a secure private key. G has three down sampled, six residual, two transposed, and one convolutional layer. Three times down sampling picture characteristics. Six identical blocks [37] make things. Then, two convolutional layers are transposed. Convolution makes pictures from low-level features. Outputs DeepKeyGen's private key. Instance normalization enhances picture quality, speeds model convergence, and avoids gradient explosion.

$L_G = \min(E_{x\sim pdata(x)} \log(1 - D(G(x)))$ (2)

x represents the beginning image, The created key is near the transformation domain, so the discriminator believes it's from there.

2) Discriminator Network D: Checks whether picture is in transformation domain. Five difficult stages in Dis. Four convolution layers extract features. Last layer processes 1-D characteristics. This identifies fakes. Network D's accuracy is 50% when the networks balance. D can't distinguish between the produced key and transformation domain key.

### 3.1 Private Key Image

DeepKeyGen's private key is a stream cypher picture. Pixels make up each picture. These pixels include value and geographical information. The private key is thus a composite of picture pixels.

$KEY_{definition} = [V_1, V_2, \ldots V_i, \ldots V_n]$ (3)

In the equation, Vi is one pixel.It's a key sequence value.

SEYBOLD Report
ISSN: 1533 - 9211
DOI: 10.5281/zenodo.7135160

$V_i = [p_i, x, y, c]$ (4)

pi is the pixel value, x is vertical, and y is horizontal. From 0 to 255, Pi, x, y, and c range. 4D private key, not stream cyphers. The key values (pixel values and 3-D space location information) make the private key challenging and increase its security. Networks learn source-to-transformation mapping from keys. Before training, convolutional parameters are random. I indicates the ith DeepKeyGen convolution layer. DeepKeyGen W is convolutional.

W=[W1,W2.. .Wn] (5)
W=x=G (6)

W indicates network settings and x is the initial image. Convolutional networks turn input images into feature vectors during training. Forward propagation produces the original private key, which is used to compare the current and target private keys in the transformation domain. Backpropagation transfers loss to convolutional layers. Back propagation using gradient descent improves performance

(Wjn, I = Wj1n, I J (Wjn)) (7)

J (Wjn) represents the gradient of the nth convolutional layer's Jth training loss. Gradient descent improves network mapping. G and D are distinct. After the training phase, the loss stabilizes and the transformation domain key is established. Alg. 1 depict the keygeneration procedure.

**3.2 Algorithm: Key Generation Algorithm.**

Initialization: Set the DeepKeyGen's W parameters to a random value using the formula Wn = random [wn,1, wn,2,..., wn,i]. KEY is 256× 256× 3.

Step 1: Convert training images to a 256× 256× 3 matrix. x = Convert(IMAGE source domain); y = Convert (IMAGE transformation domain).
Step2: Forward propagation of generator network G, key G (x).
Step 3: G's deepest layer produces KEY. D(y) = Result / Forward propagation of discriminator network D.
Step 4: D(y) = Result / Forward propagation of D. The transformation domain judgment should be printed.
Step 5. LG + LD = L / Determine the overall loss. It's backwards propagation, or the number It's backwards propagation, or the number and it's happening in reverse.
Step 6: Wjn,i = Wj1n,i J(Wjn) / It is necessary to compute the gradient that will be sent back to each layer before updating the network settings.

SEYBOLD Report
ISSN: 1533 - 9211
DOI: 10.5281/zenodo.7135160

Step 7: The output, KEY, is quite near to the transformation domain, so we'll just stop here and call it a day.

## 4. PROPOSED ALGORITHM

DNA cryptography and Feistel cipher improve computer security. Add the DNA D-box table to the Feistel Cipher to compute. The D-box swaps codons 43 times. Three DNA bases dictate the location of an amino acid in a protein molecule during protein synthesis. 20 amino acids (A, B, C) are made from 61 codons, and 3 are stop signals. Phenylalanine (F) may be substituted with TTT or TTC, producing confusion in letter order. In this work, amino acid names aren't supplied and 64-codons (Figure 1) are constructed randomly (8 x 8 grids). 64 combinations of 3 DNA bases from 4 create codons. Block 64-bit calculation If bits are less than 64, zeroes are added.

$E_K(M) = C$ indicates encryption, whereas $D_K(C) = M$ shows decoding. Plaintext (M) is any character; encrypted text $C \in \{A,C,G,T\}^*$. Sender and recipient must agree on the number of rounds and arbitrary keys. Random key $K \in \{0, 1\}^*$

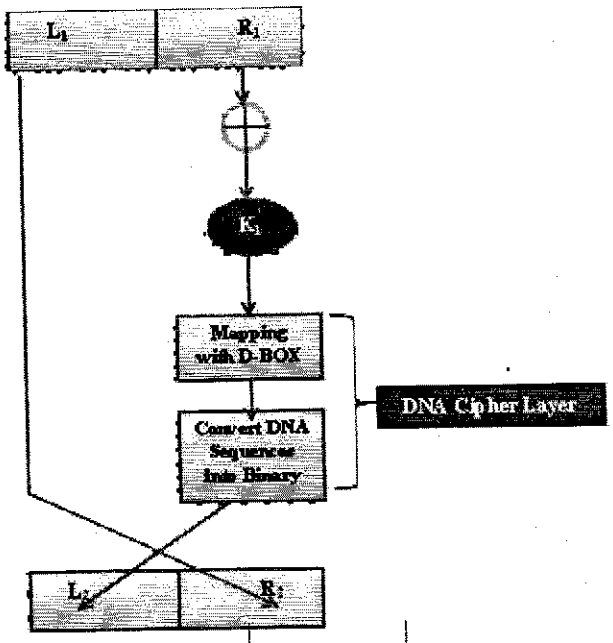



**Figure 1: Inspired Feistel DNA Cryptosystem using D-Box**

Bob needs to hear from Alice. They agree on the number of rounds, and Alice splits the plaintext into 32-bit pairs. Every round needs a unique key. XOR key and content (Ri). Add 4 zeros to get octal. This sets the D-Box codon. DNA is binary (A-00, C-01, G-10, T-11). Remove the remaining four cushioned pieces and exchange them. After n cycles, they offer Bob the binary sequence as DNA. Bob gets the cipher and keys via an open channel. Bob reverses and gets plaintext.

SEYBOLD Report
ISSN: 1533 - 9211
DOI: 10.5281/zenodo.7135160

| D-BOX | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|---|---|
| 0 | TTT | TTC | TCT | TTG | TTA | TCC | TCA | TCG |
| 1 | TAT | TAC | TGT | TAG | TAA | TGC | TGA | TGG |
| 2 | ATT | ATC | ACT | ATG | ATA | ACC | ACA | ACG |
| 3 | CAT | CAC | CGT | CAG | CAA | CGC | CGA | CGG |
| 4 | CTT | CTC | CCT | CTG | CTA | CCC | CCA | CCG |
| 5 | AAT | AAC | AGT | AAG | AAA | AGC | AGA | AGG |
| 6 | GTT | GTC | GCT | GTG | GTA | GCC | GCA | GCG |
| 7 | GAT | GAC | GGT | GAG | GAA | GGC | GGA | GGG |

**Figure 2: D-Box of Inspired Feistel DNA Cryptosystem**

**4.1 Algorithm for Encryption**

**Algorithm Encryption(M,num)**

**M is the plaintext and num is number of rounds to be performed**

**BEGIN**

1. Transform the Message M of into ASCII and in turn into binary values.
2. Divide the message into two equal halves names as $L_1$ and $R_1$
3. For j=1 to num

Begin

1. Randomly generate the key of length 32-bit and store it in $K_j$.
2. $R_j=R_j \oplus K_j$
3. $R_j=R_j+$"0000"
4. Rj has 36 3-bit bits. Convert to octal.
5. Consider pairwise octal values and map them into D-box then the results is 6 codon values.
6. Transform the DNA sequence into its equivalent binary.
7. Delete the last four padded characters and store it in $R_i$.
8. Swap $L_j$, $R_j$.

End

4. $C=L_j||R_j$.
5. C is ciphertext (00-A, 11-T, 10-G, 01-C)..

**END**

**4.2 Algorithm for Decryption**

**Algorithm Decryption(C, num)**

C is cipher text and num is number of rounds. K is an array of set of keys for number of rounds received from the sender through secure medium.

**BEGIN**

1. Transform the cipher text C into its equivalent binary form according DNA digital coding.

SEYBOLD Report
ISSN: 1533 - 9211
DOI: 10.5281/zenodo.7135160

2. Divide C into 32-bit Lnum and Rnum..
3. For j=num to 1

Begin

i) Lj=Lj+"1111"
ii) Convert binary to DNA bases (00-A, 11-T, 10-G, 01-C).
iii) Codonize DNA sequences.
iv) Retrieve D-Box codon locations.
v) Convert octal locations to binary
vi) Remove "0000" from LJ.
vii) Lj=Lj□Kj
viii) Swap Lj&Rj.

End

4. $M=L_j||R_j$
5. Convert the binary data to ASCII M.

**END**

### 4.3 Empirical Analysis

#### 4.3.1Process of Encryption

Let us take the plaintext as M=Grooming and number rounds n=4. Convert the plaintext into ASCII and in turn into binary then,

| Character | ASCII | Binary |
|---|---|---|
| G | 71 | 01000111 |
| r | 114 | 01110010 |
| o | 111 | 01101111 |
| o | 111 | 01101111 |
| m | 109 | 01101101 |
| i | 105 | 01101001 |
| n | 110 | 01101110 |
| g | 103 | 01100111 |

**First Round**

a) Perform XOR operation between right part of previously obtained result with randomly generated Key value.

$R_1$ =01101101011010010110111001100111
$K_1$ =10010000001101011100011011011000
$R_1 \oplus K_1$ =11111101010111001010100010111111

b) Divide the binary into octal values. The result of XOR contains only 32 bits when we split them 3 bits each and make a pair we need to pad four more bits to the result.

111,111,010,101,110,010,101,000,101,111,110,000
7,7,2,5,6,2,5,0,5,7,6,0

c) Retrieve the codon value which placed in the D-box.

GGG,ACC,GCT,AAT,AGG,GTT

d) Convert them into equivalent binary values according to DNA Digital Coding.

10101000010110011100001100101010111 1

e) Remove final four bits and swap $L_1$ and $R_1$

10101000010110011100001100101010010001110111001001101111011011111

**Second Round**

a) Perform XOR operation between right part of previously obtained result with randomly generated Key value.

$R_2$ =01000111011100100110111101101111

$K_2$ =01011010110010110010010100101011

$R_2 \oplus K_2$ =00011101101110010100101001000100

b) Divide the binary into octal values. The result of XOR contains only 32 bits when we split them 3 bits each and make a pair we need to pad four more bits to the result.

000,111,011,011,100,101,001,010,010,001,000,000

0,7,3,3,4,5,1,2,2,1,0,0

c) Retrieve the codon value which placed in the D-box.

TCG,CAG,CCC,TGT,ATC,TTT

d) Convert them into equivalent binary values according to DNA Digital Coding.

11011001001001010111101100110111111 1

f) Remove final four bits and swap $L_2$ and $R_2$

11011001001001010111101100110111101010000101100111000011001010 10

**Third Round**

a) Perform XOR operation between right part of previously obtained result with randomly generated Key value.

$R_3$ =10101000010110011100001100101010

$K_3$ =01001011100110100011001001011101

$R_3 \oplus K_3$ =11100011110000111111000101110111

a) Divide the binary into octal values. The result of XOR contains only 32 bits when we split them 3 bits each and make a pair we need to pad four more bits to the result.

111,000,111,100,001,111,110,001,011,101,110,000

7,0,7,4,1,7,6,1,3,5,6,0

b) Retrieve the codon value which placed in the D-box.

GAT,GAA,TGG,GTC,CGC,GTT

c) Convert them into equivalent binary values according to DNA Digital Coding.

10001110000011101010110101100110111

g) Remove final four bits and swap $L_3$ and $R_3$

10001110000011101010110101100110111011001001001010111101100110111

**Fourth Round**

a) Perform XOR operation between right part of previously obtained result with randomly generated Key value.

SEYBOLD Report
DOI: 10.5281/zenodo.7135160
ISSN: 1533 - 9211

$R_4$ =11011001001001010111101100110111

$K_4$ =01101001010100010010101001010101

$R_4 \oplus K_4$ =10110000011101000101000101100010

a) Divide the binary into octal values. The result of XOR contains only 32 bits when we split them 3 bits each and make a pair we need to pad four more bits to the result.

101,100,000,111,010,001,010,001,011,000,100,000

5,4,0,7,2,1,2,1,3,0,4,0

b) Retrieve the codon value which placed in the D-box.

AAA,TCG,ATC,ATC,CAT,CTT

c) Convert them into equivalent binary values according to DNA Digital Coding.

00000011011000110100110101001011011111

d) Remove the last four bits and interchange $L_4$ and $R_4$ then the string is

000000110110001101001101010011011000111000001110101011010110011 0

After 4 rounds, the binary data is converted into DNA form which is a cipher texti.e.,

00,00,00,11,01,10,00,11,01,00,11,01,01,00,11,01,10,00,11,10,00,00,11,10,10,10,11,01,01,10, 01,10

AAATCGATCATCCATCGATGAATGGGTCCGCG

**4.3.2 Process of Decryption**

The received Cipher text from the Sender is

C= AAATCGATCATCCATCGATGAATGGGTCCGCG

1. Transform the ciphertext C into its equivalent binary form as per the DNA Digital Coding.

000000110110001101001101010011011000111000001110101011010110011 0

**First Round**

a) Divide the ciphertext into two parts $L_4$ and $R_4$.

$L_4$=00000011011000110100110101001101

$R_4$=10001110000011101010110101100110

b) $L_4$=00000011011000110100110101001101 $L_4$=AAATCGATCATCCATC

Split them into codons and find their positions in the D-Box and pad four bits to make them into codons.

$L_4$=AAA,TCG,ATC,ATC,CAT,CTT,

i.e. In D-box, the position of AAA is 5th row and 4th column and that of TCG is 0th row and 7th column. This process is repeated for remaining codons.

L4=5,4,0,7,2,1,2,1,3,0,4,0

Convert octal to binary and eliminate padding.

$L_4$=10110000011101000101000101100010 0000

After removing the padded bits,

$L_4$=10110000011101000101000101100010

c) Perform XOR operation between $L_4$ with $K_4$

DOI: 10.5281/zenodo.7135160
ISSN: 1533 - 9211

$L_4$ = 10110000011101000101000101100010

$K_4$ = 01101001010100010010101001010101

$L_4 \oplus K_4$ = 11011001001001010111101100110111

d) Interchange $L_4$& $R_4$ and convert into DNA form

$L_4$=11011001001001010111101100110111

$R_4$=10001110000011101010110101100110111001001001010111101100110111

C=1000111000001110101011010110011011011001001001010111101100110111

**Second Round**

a) Divide the Ciphertext received from the previous round into two equal parts names as $L_3$ and $R_3$.

$L_3$=10001110000011101010110101100110

$R_3$=11011001001001010111101100110111

b) Take the left part of the data

$L_3$= 10001110000011101010110101100110

Convert them into DNA and then split them into codons and pad four bits to make even number of codons.

GAT, GAA, TGG, GTC, CGC, GTT

Find the position of codons from the D-Box.

$L_3$=7,0,7,4,1,7,,6,1,3,5,6,0

c) Convert the octal values into binary and remove the padded four bits.

$L_3$=111000111100001111111000101110110000

After removing padded bits

$L_3$=11100011110000111111000101110111

d) Perform XOR operation between $L_3$ with $K_3$

$L_3$ =11100011110000111111000101110111

$K_3$ =01001011100110100011001001011101

$L_3 \oplus K_3$ =10101000010110011100001100101010

e) Interchange $L_3$& $R_3$ and convert into DNA form

$L_3$=10101000010110011100001100101010

$R_3$=11011001001001010111101100110111

C=1101100100100101011110110011011110101000010110011100001100101010

**Third Round**

a) Divide the ciphertext of the previous round into two equal parts i.e., $L_2$ and $R_2$

$L_2$=11011001001001010111101100110111

$R_2$=10101000010110011100001100101010

b) Take the left part

$L_2$=11011001001001010111101100110111

Convert them into DNA, split them into codons, and pad four bits to make them even codons

$L_2$=TCG,CAG,CCC,TGT,ATC,TTT

SEYBOLD Report
ISSN: 1533 - 9211
DOI: 10.5281/zenodo.7135160

Find the positions of each codon from the D-Box

$L_2$=0,7,3,3,4,5,1,2,2,1,0,0

c) Convert the octal values into binary and remove the padded bits.

$L_2$=000111011011100101001010010001000000

After removing padded bits

$L_2$=000111011011100101001010010001000100

d) Perform XOR operation between L2 with K2.

$L_2$ =000111011011100101001010010001000100

$K_2$ =010110101100101100100101001010011

$L_2+\oplus K_2$ =010001110111001001101111011011111

e) Interchange $L_2$& $R_2$ and convert them into DNA

$L_2$=010001110111001001101111011011111

$R_2$=101010000101100111000011001010101

C=101010000101100111000011001010100100011101110010011011110110111111

**Fourth Round**

a) Divide ciphertext into two equal parts i.e., $L_1$ and $R_1$.

$L_1$=101010000101100111000011001010101

$R_1$=010001110111001001101111011011111

b) Take the left part and convert into DNA form

$L_1$=101010000101100111000011001010101

$L_1$=GGGACCGCTAATAGGG

Split them into Codons and pad four bits into binary form to make them even codons

$L_1$=GGG,ACC,GCT,AAT,AGG,GTT

Find the positions codons from the D-Box

$L_1$=7,7,2,5,6,2,5,0,5,7,6,0

c) Convert the octal values into binary form and remove the padded bits.

$L_1$=111111010101110010101000101111110000

After removing the padded bits,

$L_1$=111111010101110010101000010111111

d) Perform XOR operation with $L_1$& $K_1$.

$L_1$ =111111010101110010101000010111111

$K_1$ =100100000011010111000110110110000

$L_1\oplus K_1$ =011011010110100101101110011001111

e) Interchange $L_1$ and $R_1$and convert them into DNA.

$L_1$=011011010110100101101110011001111

$R_1$=010001110111001001101111011011111

C=010001110111001001101111011011110110110101101001011011100110011 1

C=CACTCTAGCGTTCGTTCGTCCGGCCGTGCGCT

At final round, the cipher text convert into binary and then into equivalent ASCII character

SEYBOLD Report
ISSN: 1533 - 9211
DOI: 10.5281/zenodo.7135160

C=CACTCTAGCGTTCGTTCGTCCGGCCGTGCGCT

C=01000111011100100110111101101111011011010110100101101110011001111

| Binary Data | ASCII Value | ASCII Character |
|---|---|---|
| 01000111 | 73 | G |
| 01110010 | 114 | r |
| 01101111 | 111 | o |
| 01101111 | 111 | o |
| 01101101 | 109 | m |
| 01101001 | 105 | i |
| 01101001 | 110 | n |
| 01100111 | 103 | g |

## 5. RESULTS

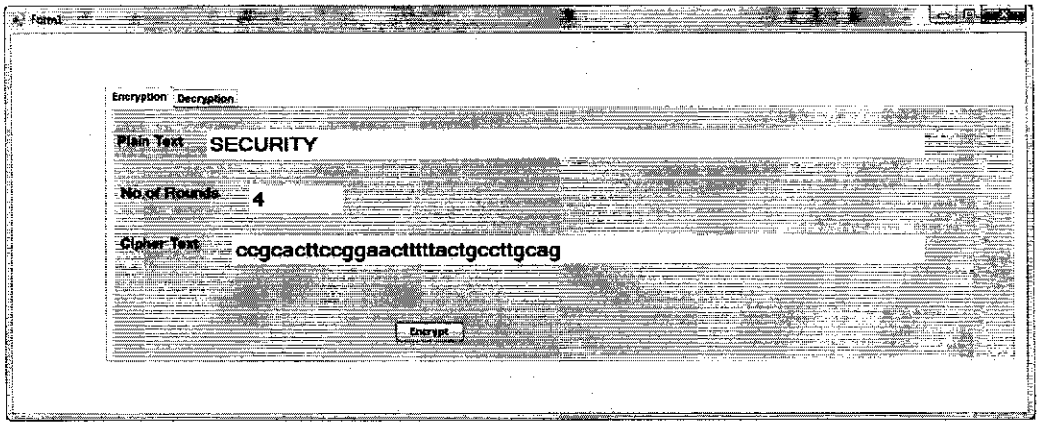
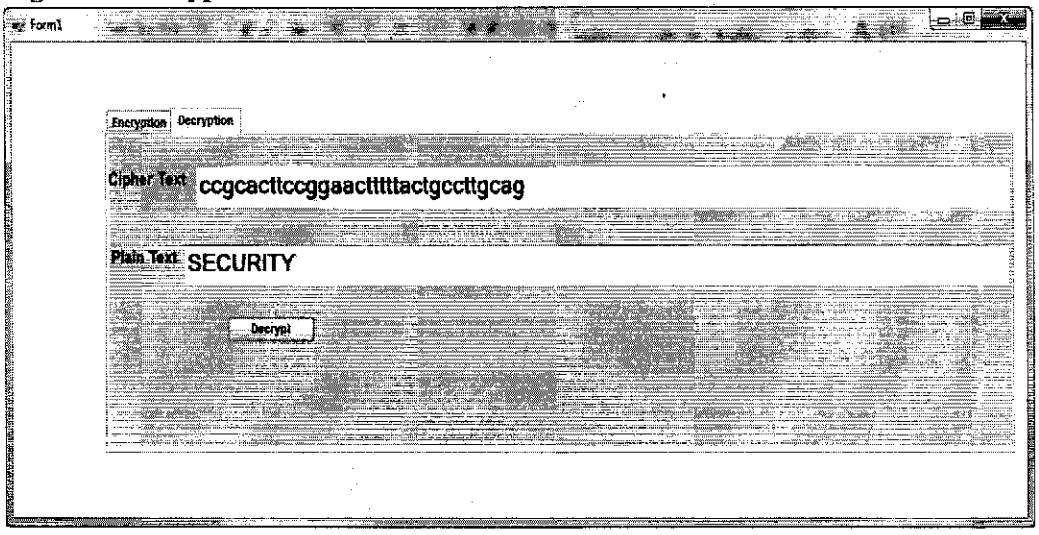



**Figure 3: Encryption Process for the Plaintext "SECURITY" for 4 rounds.**



SEYBOLD Report
ISSN: 1533 - 9211
DOI: 10.5281/zenodo.7135160

**Figure 4: Decryption process for the cipher text "ccgcacttccggaacttttttactgccttgcag".**

## 5.1 PERFORMANCE MEASUREMENT

The analysis of an algorithm measured in Intel i3 processor in .NET environment for the string "SECURITY" with 4 number of rounds. The following tables showed the difference in the time taken for encryption and decryption process between the proposed model and the existing traditional feistel approach model .It is observed that, though the proposed model takes more time because of DNA cipher layer, still it may be preferred because of more security.

**Table 1 : Comparison table of Encryption time between Proposed and Traditional approach**

| Number of Rounds | Encryption Time for Proposed Algorithm | Encryption Time for Traditional Approach |
|---|---|---|
| 4 | 0.011 | 0.00423 |
| 5 | 0.012 | 0.00456 |
| 6 | 0.013 | 0.00487 |
| 7 | 0.013 | 0.00512 |
| 8 | 0.0145 | 0.00543 |
| 9 | 0.0154 | 0.00654 |
| 10 | 0.018 | 0.00543 |

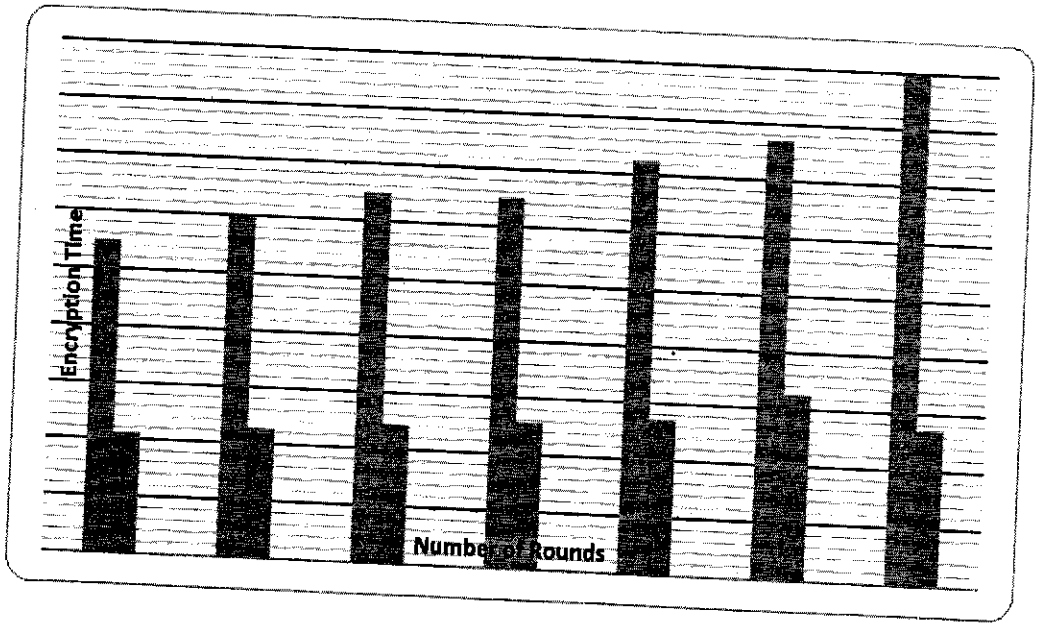



**Figure 5: Performance analysis for the message SECURITY in relation between number of rounds and encryption time.**

SEYBOLD Report
ISSN: 1533 - 9211
DOI: 10.5281/zenodo.7135160

**Table 2: Comparison of Decryption time between Proposed and Traditional Approach**

| No.of Rounds | Decryption Time for Proposed Algorithm | Decryption Time for Traditional Approach |
|---|---|---|
| 4 | 0.0032 | 0.0018 |
| 5 | 0.0048 | 0.0017 |
| 6 | 0.0064 | 0.0023 |
| 7 | 0.0089 | 0.0045 |
| 8 | 0.0065 | 0.0035 |
| 9 | 0.011 | 0.0034 |
| 10 | 0.0123 | 0.0032 |

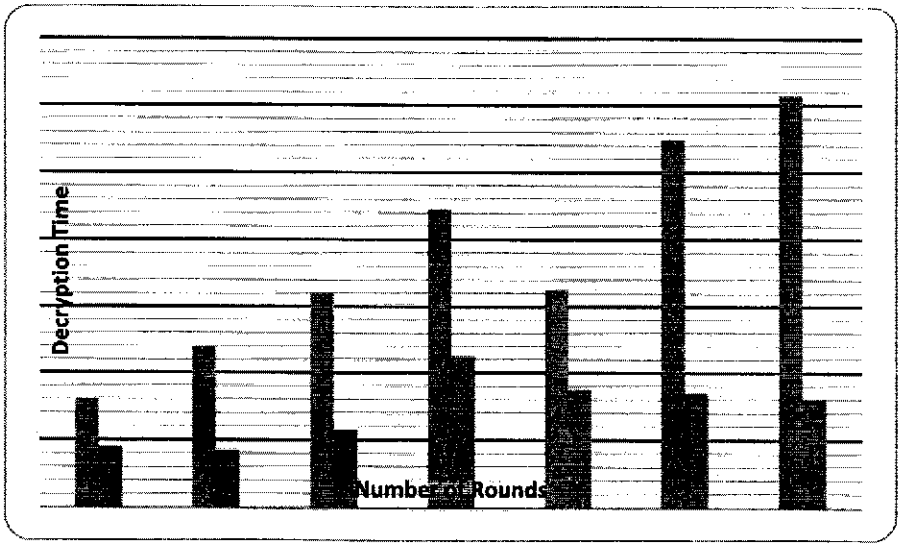



**Figure 6: Performance analysis for the message SECURITY in relation with number of rounds and time.**

## 5.2 AVALANCHE EFFECT

The following table 5.3 represents the percentage of bits flipped in the cipher text when a single character in the plaintext is modified.

Plaintext1= security

Plaintext2=secxrity

Total 34 bits are flipped in the cipher text when change of single character in the plaintext are shown in red colour.

Ciphertext1=

01110011011001010110001101111000011100100110100101110100011111001

SEYBOLD Report
ISSN: 1533 - 9211
DOI: 10.5281/zenodo.7135160

Ciphertext2=
00111100110100110001000110010000101101100011000111011010000000010

**Table 3:Comparison of Avalanche Effect for strings security and secxrity to the proposed and traditional models**

| Name of the Algorithm | Number of bits flipped | % |
|---|---|---|
| Ceaser Cipher | 2 | 4.132 |
| Playfair | 3 | 4.76 |
| Vigener Cipher | 2 | 4.124 |
| Blowfish | 19 | 25.78 |
| DES | 32 | 49.65 |
| Feistel Inspired DNA based Cryptosystem | **34** | **54.13** |

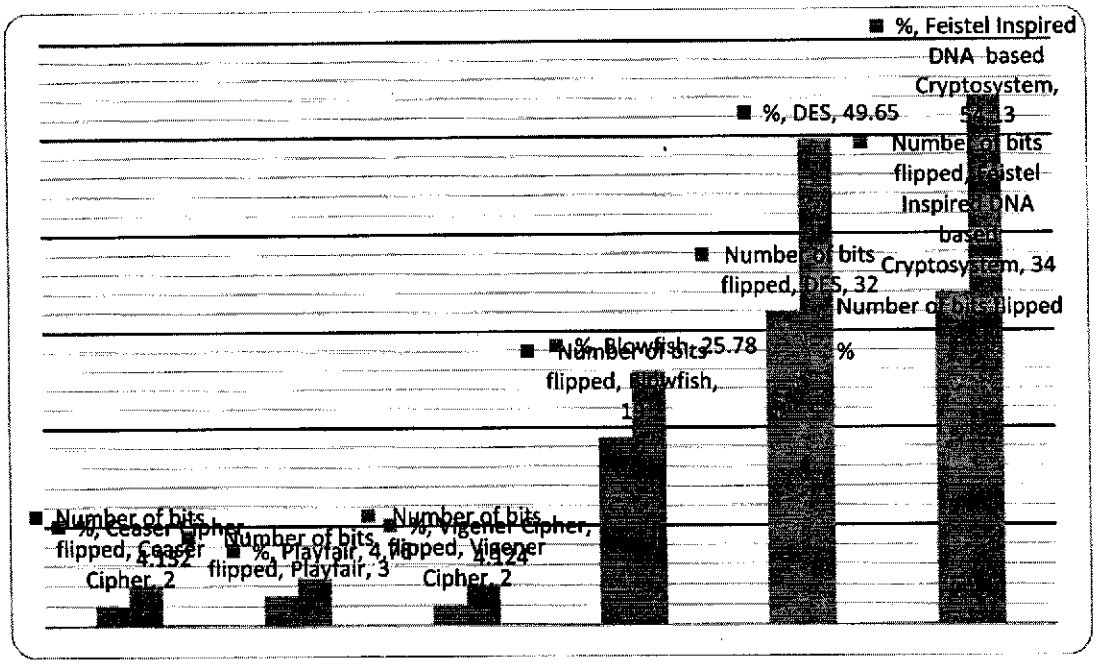



**Figure 7 :Comparison of the results of the avalanche effect for strings security and secxrity.**

In the considered plaintext "security", the change of one character leads to change in 34 bits in the ciphertext which was very difficult to the intruder in identifying the plaintext. It was observed that the total number of bits flipped in the present model is more when compared to existing models.

SEYBOLD Report
ISSN: 1533 - 9211
DOI: 10.5281/zenodo.7135160

## 6. CONCLUSION & FUTURE SCOPE

In this, DNA codons are organized in a D-box for a symmetric blockcipher. 64-bit plaintext is split in two, the randomly generated key is XORed with the right segment, and the D-codon box values are obtained. DNA's binary values encode the text. Encryption text is separated at the receiver. The left section becomes DNA bases and codons. The D-Box codons are converted to octal and XORed with the key value. In each cycle of 64-bit plaintext ciphering, most bits change. Simple text can't be hacked. Compares two models' performance. Avalanche effect is explored. Even though DNA is cited as a medium for ultra-minimal data storage and the future seems bright, there are still worries like quantum attacks and bio molecular labs. Future work may include building a public key cryptosystem using PCR Amplification. Increasing LOOKUP table size for additional character sets improves performance. More compression methods will be used on DNA sequences to improve compression ratio, and combining encryption and compression leads to a stronger cryptosystem. Future assaults will be practiced. More models will be added by increasing Avalanche effect.

**References:**

[1] A. S. Abad, H. Hamidi, "An Architecture for Security and Protection of Big Data", in International Journal of Engineering (IJE), TRANSACTIONS A: Basics Vol. 30, No. 10, (October 2017), pp. 1479-1486

[2] B. B. Raj, J. Frank, T.Mahalakshmi, "Secure Data Transfer through DNA Cryptography using Symmetric Algorithm", in International Journal of Computer Applications, Vol 133-No 2, pp. 0975-8887, January 2016

[3] A. Roy, A. Nath, "DNA Encryption Algorithms: Scope and Challenges in Symmetric Key Cryptography", in International Journal of Innovative Research in Advanced Engineering, ISSN: 2349-2763, Issue 11, Volume 3, Nov, 2016.

[4] N. S. Kolte, K. V. Kulhalli and S. C Shinde, "DNA Cryptography using Index-based Symmetric DNA Encryption Algorithm", International Journal Of Engineering Research and Technology, ISSN 0974-3154 Vol 10, No1 ,2017.

[5] S. Karthiga, E. Murugavalli, "DNA Cryptography", in International Research Journal of Engineering and Technology, p-ISSN 2395-0072, Vol 5, March 2018.

[6]. Bonny B.Raj, V. Ceronmani sharmimila," An Survey on DNA Based Cryptography" IEEE 2018 International Conference on Emerging Trends and Innovations In Engineering And Technological Research (ICETIETR) - Ernakulam (2018.7.11-2018.7.13)] 2018.

[7]. Saijisha K S,S.Mathew," An encryption based on DNA cryptography and steganography"IEEE 2017 International conference of Electronics, Communication and Aerospace Technology (ICECA)COIMBATORE, India (2017.4.20-2017.4.22)] 2017 .

[8]. S.Roy, Sudipta Singha, Shahriyar, Shaikh Akib, Asaf-Uddowla, Md, Alam, Kazi Md. Rokibul; Morimoto, Yasuhiko"A novel encryption model for text messages using delayed chaotic neural network and DNA cryptography[IEEE 2017 20th International Conference of Computer and Information Technology (ICCIT) - Dhaka, Bangladesh(2017.12.22-2017.12.24)].

SEYBOLD Report
ISSN: 1533 - 9211
DOI: 10.5281/zenodo.7135160
[9]. K.KALAISELVI "Enhanced AES Cryptosystem by using Genetic Algorithm and Neural Network in S-box 978-1-5090-1936-6/16/$31.00 ©2016 IEEE. 5. Panagiotis Papadimitratos". Secure Data Communication in Mobile Ad Hoc Networks", IEEE journal on selected areas in communications 0733-8716.

[10]. Md .Rafiul Biswas,Kazi Md.Rokibul Alam, Ali Akber,Ya Suhiko Mori-moto"A DNA cryptographic technique based on dynamic DNA encoding and asymmetric cryptosystem" Published in 2017 4th International Conference on networking system and security(NSysS)IEEE.

[11]A. Nath, A. Dodia, "Symmetric Key Encryption Algorithm using DNA Sequence," Int. J. Adv. Res.Comput. Sci. Manag. Stud., vol. 6, no. 4, pp. 108–115, 2018.

[12] K. C. Jithin, S. Sankar, "Colour image encryption algorithm combining Arnold map, DNA sequence operation, and a Mandelbrot set," J. Inf. Secur. Appl., vol. 50, no. 102428, pp. 1–22, Feb. 2020.

[13] A. Hazra, C. Lenka, A. Jha, M. Younus, "A Novel. Two Layer Encryption Algorithm Using One-Time Pad and DNA Cryptography," in Lecture Notes in Networks and Systems, vol. 103, 2020, pp. 297–309.

[14] P. B. Narasingapuram , M. Ponnavaikko, "DNA Cryptography Based User Level Security forCloud Computing and Applications," Int. J. Recent Technol. Eng., vol. 8, no. 5, pp. 3738–3745, 2020.

[15] A. Kaushik, V. Thada, "VG1 Cipher – A DNA Indexing Cipher," Int. J. Innov. Technol. Explor.Eng., vol. 9, no. 3, pp. 221–226, 2020.

[16] Rajni, "DNA Computing," Int. J. Eng. Comput. Sci., vol. 6, no. 1, pp. 19972–19976, 2017.

[17] O. Tornea, M. E. Borda, V. Pileczki,"Cryptographic Algorithm Based on Dna and RNA Properties," Int. J. Adv. Res. Comput. Eng. Technol., vol. 7, no. 11, pp. 237–241, 2018.

JOURNAL OF OPTOELECTRONICS LASER
Volume 41 Issue 11, 2022
ISSN: 1005-0086

# Deep Learning-Based Key Generation for DNA ASCII Table-Based Encryption

**G. Ramaswamy*[1] & Dr. R. Satya Prasad2**
*[1]Research Scholar, Acharya Nagarjuna University
[2]Professor in the Department of Computer Science and Engineering, Acharya Nagarjuna University

**ABSTRACT**
New encryption methods are required to safeguard data during transmission over a network or while stored in the cloud, since many existing cryptography algorithms have been cracked. DNA Cryptography been developed in this context to ensure the safety of data stored in DNA sequences. DNA cryptography methods are implemented using a wide variety of bio molecular approaches. DNA's inherent randomness has rendered many encryption techniques useless. Researchers have introduced additional cryptographic methods due to DNA's capacity to store massive amounts of data (almost peta bytes) in very little amounts of the molecule. A new cryptographic technique is proposed in this study, splitting the message in half and then XORing each half with a random value.

**Keywords:** Encryption, Decryption, DNA Cryptography, SPRIAL MODEL

## I. INTRODUCTION

The protection of sensitive information is a concern for every company. Threats have also increased along with the development of technology. Any company must place a high priority on ensuring the confidentiality of its customers' data while it is being sent. The information can only be accessed by the sender and the recipient, since this is the most important requirement. As a consequence of this, the researchers expressed an interest in developing a variety of different security techniques. Utilizing cryptographic procedures is one method for ensuring the data's safety. [Closed captioning] [Open captioning] The data may be sent over the internet in a safe manner when these methods are used. At the sender's end, the information is sent in an unreadable format via a process known as encryption. At the receiver's end, however, the unreadable format may be turned into a readable one through a process known as decryption [1]. The format that cannot be read.

This paper introduces a table called DNA ASCII Table which contains 256 DNA Strands that in turn contains 4 bases. These strands mapped with the original ASCII table. In the third chapter, LOOKUP table that contains only 64 values was discussed. The LOOKUP table alphabet set contains Uppercase English alphabets, Lowercase English alphabets, Digits from 0..9, Special characters fullstop and single space. All these LOOKUP table alphabets are mapped with DNA Codons and from the observations it was clear that it is limited to 64 values because the LOOKUP table permits 64! Permutations to find the cipher text. This drawback can be overcome by proposing DNA ASCII Table which allows the permutations up to 256! In each small DNA Strand Four DNA bases can be placed in 4*4*4*4=256 distinctive sequences.

DNA cryptography techniques are outlined in detail in Section 2, whereas Section 3 details the proposed algorithmic approaches. Experiment findings and analysis of the effectiveness of the proposed technique are presented in Section 4, and Section 5 ties things up by analyzing the conclusions and the prospective expansions of the study.

## II. RELATED WORK

The goal is to restrict the activities of harmful users in cloud computing infrastructure. In order to generate a secure key for the user and data encryption and decryption procedure, the author recommends DNA cryptography [10]. Data from Cloud users is converted to a type of human deoxyribonucleic acid that may be used as the basis for strong encryption keys and data storage. The method is validated by its implementation in the DOTNET framework, and its experimental outcomes are confirmed.

The authors of this piece propose a solid framework for protecting intellectual property rights, ensuring that only authorized individuals may access and make unauthorized copies of data belonging to end users. The Advanced Encryption Standard algorithm and the Elliptic Curve cryptosystem are the two sound cryptographic building blocks of the framework [11] developed for Digital Right Management.

To make accessible a system that protects communication and limits illegal access, a proactive security framework [13] is sold at a garage sale. Without relying on the cloud provider's reputation, this security framework enables cloud end users to safely manage privacy and integrity of data, as well as consent to security, confidentiality, network use, and storage in the cloud. Using the AES algorithm provides a solid foundation that safeguards cloud-based information and allows access to data only when appropriate authentication and verification have taken place. The framework makes the arrangement of computational cloud services more secure, more efficient, and with less latency.


DOI: 10050086.2022.11.24

The cryptographic system plays a crucial role in protecting cloud-based information. In order to safeguard information in the cloud, a standardized encryption and decryption system must be implemented, with the key serving as the essential component. Every cloud service uses its own set of security measures to safeguard the master password. Even if the service provider always has complete access to all data and the master key, the customer cannot put their whole trust in them. In this research [14], we provide a novel approach that eliminates the need for direct key-to-hostler interaction, as well as a framework for disseminating files securely in the cloud by use of an asymmetric key and a trusted intermediary.

In order to provide a comprehensive analysis of the current software security challenges in the Cloud environment, the authors [15] have analyzed the effects of existing methods and strategies. These perspectives and the survey data inform the development of a theoretical framework for planning feasible current solutions to software security concerns in the Cloud and for presenting a preferred software securitymethod for probing the Cloud research community. Fuzzy system principles have the potential to be applied to many other aspects of Cloud security framework, making for a more robust system overall. A safe method of keeping information secret. Thus, the approach [16] combines vertical data fragmentation with multiplicative homomorphic encryption methods. Our method was put through its paces, with delays in communication, cryptography, and query processing all being simulated.

## III. PROPOSED METHOD

In this developed model also the key generated randomly which is of the length n*8 where n is the number of characters in the plaintext. This key is also transmitted securely to the receiver. Divide the plaintext into two equivalent parts and name them as $M_L$ and $M_R$ and also divide the key value into two equivalent parts and name them as $K_L$ and $K_R$. Apply cross XOR function in between the parts that is $M_L$ and $K_R$ and $M_R$ and $K_L$. Regularly the XOR operation is performed in between the right part of key and the right part of plaintext. In this model, the attacker simply know the half bits of the plaintext. But in this model cross XOR function is performed so that it is very difficult to know the bits of cipher text. After performing XOR operation convert the data into DNA bases(11-T,00-A,01-C,10-G). Split the entire DNA strand into small DNA strand which is of length four. Now, map these small DNA strands into DNA ASCII Table and replace them with their equivalent ASCII Values. After that convert these values into binary and then into DNA. Now, arrange these DNA sequences into Spiral pattern(Figure 1).Fix the column size as four and divide the DNA sequence length by four then it gives the row size. Arrange all the DNA nucleotides in the sequence in the spiral matrix. The concatenated row wise data is the final cipher text.

| | | | |
|---|---|---|---|
| 1 | 2 | 3 | 4 |
| 12 | 13 | 14 | 5 |
| 11 | 16 | 15 | 6 |
| 10 | 9 | 8 | 7 |

**Figure 1: Spiral Pattern**

PRINCIPAL
MALINENI LAKSHMAIAH
WOMEN'S ENGINEERING COLLEGE
PULLADIGUNTA, GUNTUR-17.


DOI: 10050086.2022.11.24

ISSN: 1005-0086

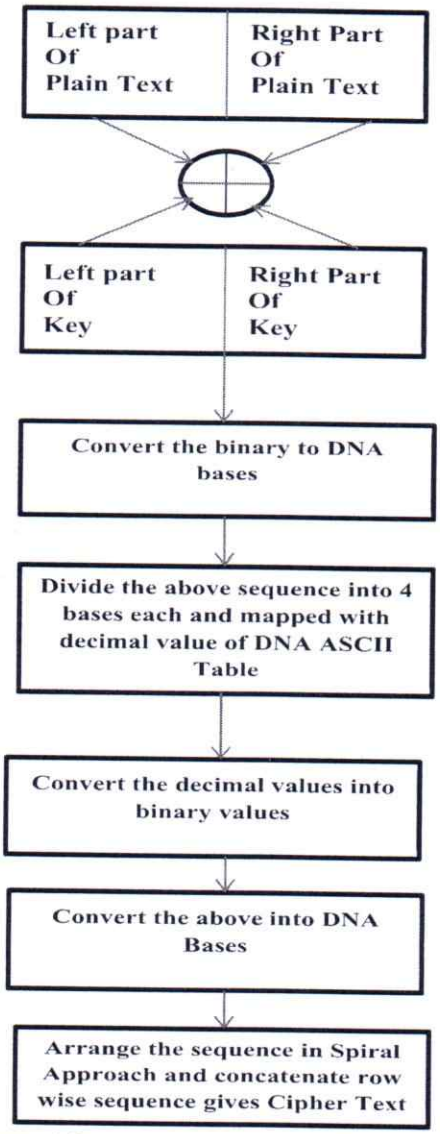



**Figure 2: DNA Based Cryptosystem using DNA ASCII Table with spiral approach**

**Algorithm for Encryption**

**Algorithm DNAASCIIEncrypt(P)**
P is the Plaintext.
**BEGIN**

1. Transform the Plaintext P into its equivalent Binary values.
2. The binary form of data is divided into $P_L$ and $P_R$.
3. Generate the randomly key such that its length should be always equal to the length of message length which is in binary form.
4. Split the key into two equal parts $K_L$ and $K_R$
5. Perform XOR between $P_L$ and $K_R$($P_L=P_L \oplus K_R$)
6. Perform XOR between $P_R$ and $K_L$($P_R=P_R \oplus K_L$)
7. Plaintext is obtained after concatenating $P_L$ and $P_R$($P=P_L+P_R$).
8. Transform P into equivalent DNA bases(00-A,11-T, G-10,01-C).
9. Split plaintext P into four bases and represent with a decimal value of DNA ASCII Table(Table 6.1).
10. Transform the decimal values into binary values and called as a variable C.


DOI: 10050086.2022.11.24

ISSN: 1005-0086

11. Transform the value of C into DNA bases (00-A,11-T,10-G,01-C).
12. Organize the DNA bases in spiral pattern (Figure 6.1) and then row wise concatenation is performed to get the cipher text.

**END**

**Algorithm for Decryption**

**Algorithm DNAASCIIDecrypt(C)**

C is the Ciphertext.

**BEGIN**

1. Arrange the cipher data in row wise manner.
2. Read the cipher text in a spiral pattern.
3. Transform DNA bases into binary values and in turn into decimal values.
4. Map the decimal values in DNA ASCII Table and retrieve the equivalent four DNA bases from it.
5. Transform DNA bases into binary values.
6. Perform XOR between $C_L$ and $K_R$ ($C_L=C_L \oplus K_R$)
7. Perform XOR between $C_R$ and $K_L$ ($C_R=C_R \oplus K_L$)
8. Cipher text is obtained after concatenating $C_L$ and $C_R$( $C=C_L+C_R$)
9. Transform C into its equivalent ASCII Character which is a true Plaintext.

**END**

## Empirical Analysis

### *Process of Encryption*

Let Plaintext M=Encryption

1. Convert the plaintext into ASCII and then into binary as follows, split them two parts named as $M_L$ and $M_R$. If it is not even, pad one character to make it as even.

E→ 69→01000101
n→110→01101110
c→ 99→01100011 — ML
r→ 114→01110010
y→121→01111001

p→112→01110000
t→116→01110100
i→105→01101001 — MR
o→111→01101111
n→110→01101110

2. Generate a Random Key from the set{0,1}* of length is n*8, where n is the number of characters in the plaintext.
3. In the considered example, the total number of characters in the plaintext message is 10 then 10*8=80 bits data generated randomly, which is the Key value.
   K=10010111101111110101010110011001111101111000100010010110010110010110100111000110
4. Split the Key value into two parts and named as $K_L$ and $K_R$.
   $K_L$=10010111101111110101010110011001111110111
   $K_R$=10001000100101100101100101101001110001 10
5. Perform cross XOR that is $M_L \oplus K_R$ and $M_R \oplus K_L$

PRINCIPAL
MALINENI LAKSHMAIAH
WOMEN'S ENGINEERING COLLEGE
PULLADIGUNTA, GUNTUR-17.


DOI: 10050086.2022.11.24

ISSN: 1005-0086

a) $M_L$=01000101011011100110001101110010011111001
$K_R$=10001000100101100101100101101001110000110
$M_L$=11001101111110000011101000011011101111111
b) $M_R$=01110000011101000110100101101111011011110
$K_L$=10010111101111110101010110011001111110111
$M_R$=11100111110010110011110011111011010011001

6. Convert the above binary form into DNA form.
M=11001101111110000011101000011011101111111111001111100101100111100111101101001100 1
M=TATCTTGAATGGACGTGTTTTGCTTAGTATTATTCGGCGC
7. Divide the plaintext into four bases and map with DNA ASCII Table.
M=TATC,TTGA,ATGG,ACGT,GTTT,TGCT,TAGT,ATTA,TTCG,GCGC
M=33,14,143,156,192,52,44,130,7,221
8. Convert the values into binary and then into DNA form
33→00100001→AGAC
14→00001110→AATG
143→10001111→GATT
156→10011100→GCTA
192→11000000→TAAA
52→00110100→ATCA
44→00101100→AGTA
130→10000010→GAAG
7→00000111→AACT
221→11011101→TCTC
9. Arrange the above data in above said spiral approach(Figure 6.1)

| | | | |
|---|---|---|---|
| A | G | A | C |
| A | A | G | A |
| C | C | T | A |
| T | T | A | T |
| A | C | G | G |
| A | T | A | G |
| A | T | A | A |
| A | C | G | T |
| T | A | A | T |
| A | T | C | G |

10. Concatenate all row data, is the final cipher text
AGACAAGACCTATTATACGGATAGATAAACGTTAATATCG

***Process of Decryption***

The cipher text from the sender is C= AGACAAGACCTATTATACGGATAGATAAACGTTAATATCG

1. Arrange the data in matrix from

| | | | |
|---|---|---|---|
| A | G | A | C |
| A | A | G | A |
| C | C | T | A |
| T | T | A | T |
| A | C | G | G |
| A | T | A | G |
| A | T | A | A |
| A | C | G | T |
| T | A | A | T |
| A | T | C | G |

2. Read the data from the spiral form(Figure 6.1), then the data is
AGACAATGGATTGCTATAAAATCAAGTAGAAGAACTTCTC
3. Divide them into four bases, convert into binary and then into decimal form




DOI: 10050086.2022.11.24

ISSN: 1005-0086

AGAC→00100001→33
AATG→00001110→14
GATT→10001111→143
GCTA→10011100→156
TAAA→11000000→192
ATCA→00110100→52
AGTA→00101100→44
GAAG→10000010→130
TACT→00000111→7
TCTC→11011101→221

4. Map the above values into DNA ASCII Table
   33→TATC
   14→TTGA
   143→ATGG
   156→ACGT
   192→GTTT
   52→TGCT
   44→TAGT
   130→ATTA
   7→TTCG
   221→GCGC
5. Convert the above DNA form into binary and perform XOR between $C_L$ and $K_R$ and $C_R$ and $K_L$
   TATCTTGAATGGACGTGTTTTGCTTAGTATTATTCGGCGC
   1100110111111000001110100001101110111111111001111100101100111100111101010011001
   $C_L$=11001101111110000011101000011011101111111
   $K_R$=10001000100101100101100101101001110000110
   $C_L$=01000101011011100110001101110010011111001
   $C_R$=11100111110010110011111001111011010011001
   $K_L$=10010111101111110101010110011001111110111
   $C_R$=01110000011101000110100101101111011011110
6. The final message
   C=01000101011011100110001101110010011110010111000001110100011010010110111101101110
   Convert the above values into equivalent ASCII Characters
   01000101→ 69→E
   01101110→110→n
   01100011→ 99→c
   01110010→114→r
   01111001→121→y
   01110000→112→p
   01110100→116→t
   01101001→105→i
   01101111→111→o
   01101110→110→n
   Now, the plaintext is Encryption.

## IV. RESULTS

PRINCIPAL
MALINENI LAKSHMAIAH
WOMEN'S ENGINEERING COLLEGE
PULLADIGUNTA, GUNTUR-17.


DOI: 10050086.2022.11.24

ISSN: 1005-0086

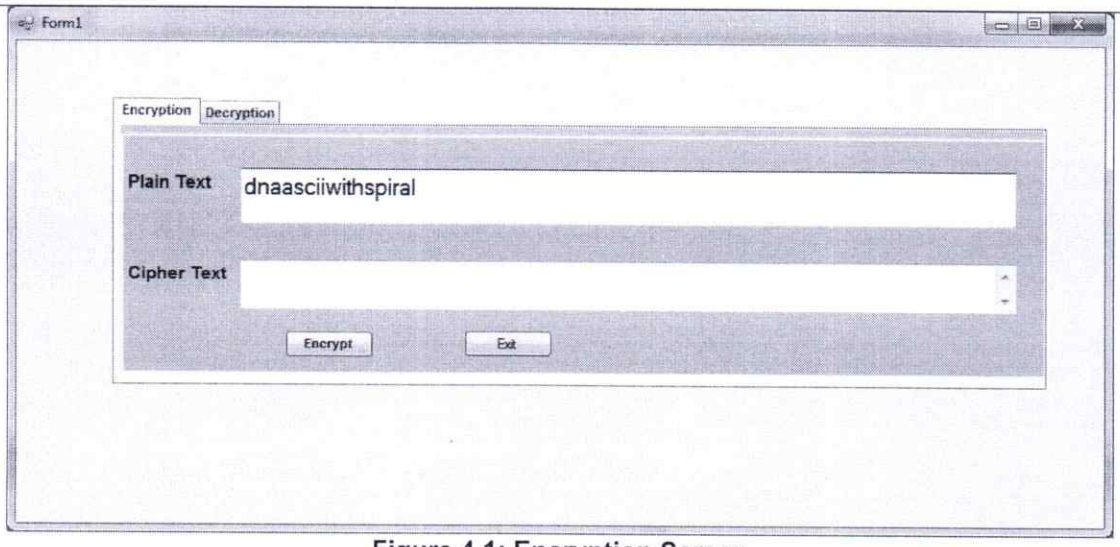



Figure 4.1: Encryption Screen

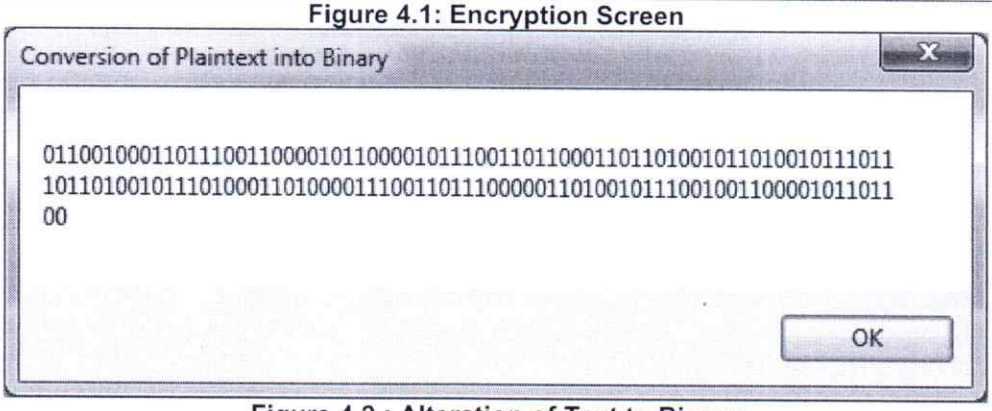



Figure 4.2 : Alteration of Text to Binary

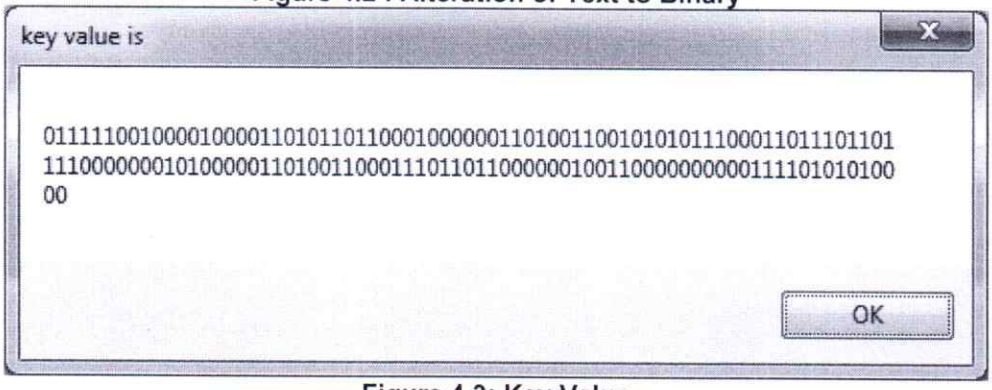



Figure 4.3: Key Value

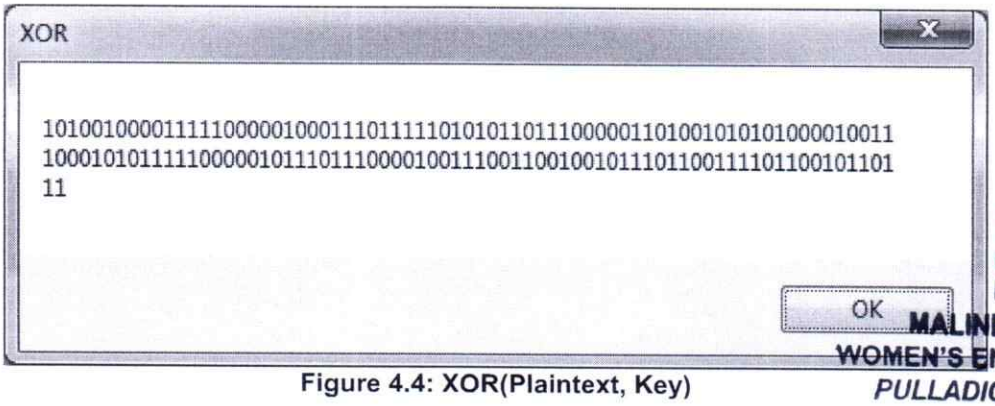



Figure 4.4: XOR(Plaintext, Key)

PRINCIPAL
MALINENI LAKSHMAIAH
WOMEN'S ENGINEERING COLL
PULLADIGUNTA, GUNTUR-1


DOI: 10050086.2022.11.24

ISSN: 1005-0086

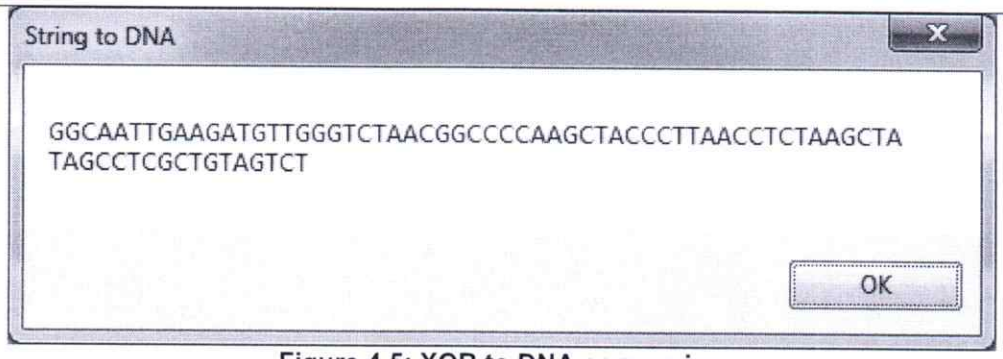



Figure 4.5: XOR to DNA conversion

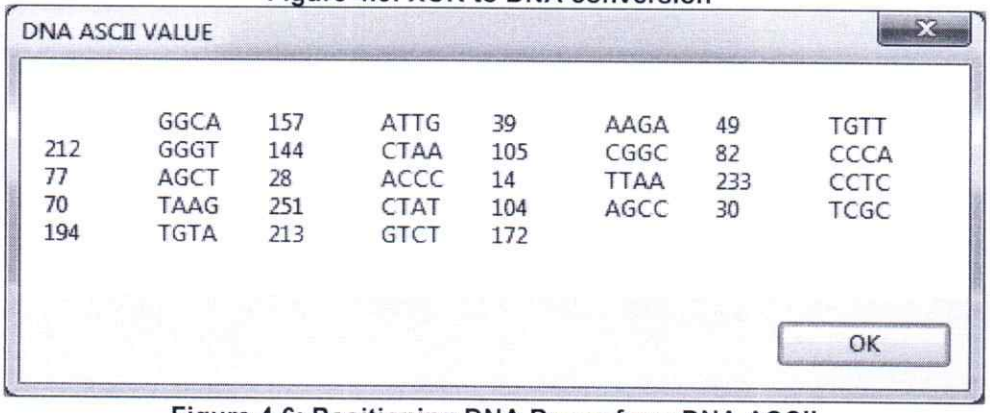



Figure 4.6: Positioning DNA Bases from DNA ASCII

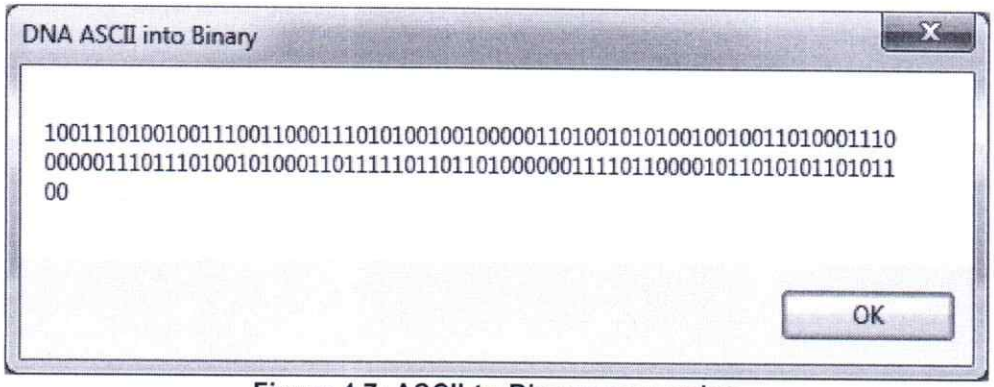



Figure 4.7: ASCII-to-Binary conversion

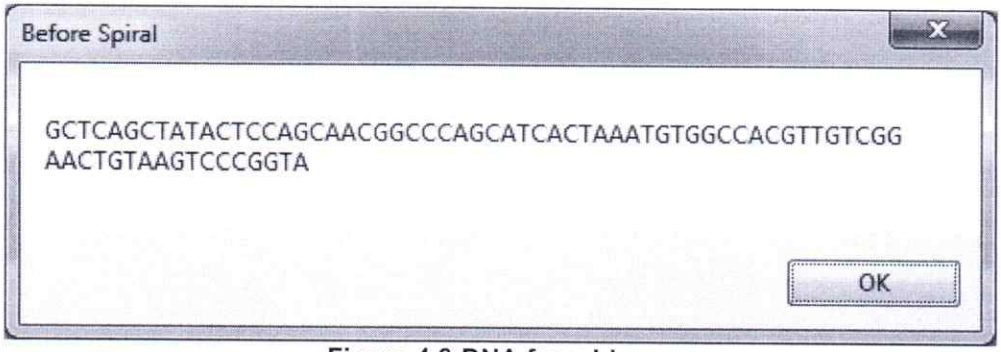



Figure 4.8:DNA from binary

PRINCIPAL
MALINENI LAKSHMAIAH
WOMEN'S ENGINEERING COLLEGE
PULLADIGUNTA, GUNTUR-17.


DOI: 10050086.2022.11.24

ISSN: 1005-0086

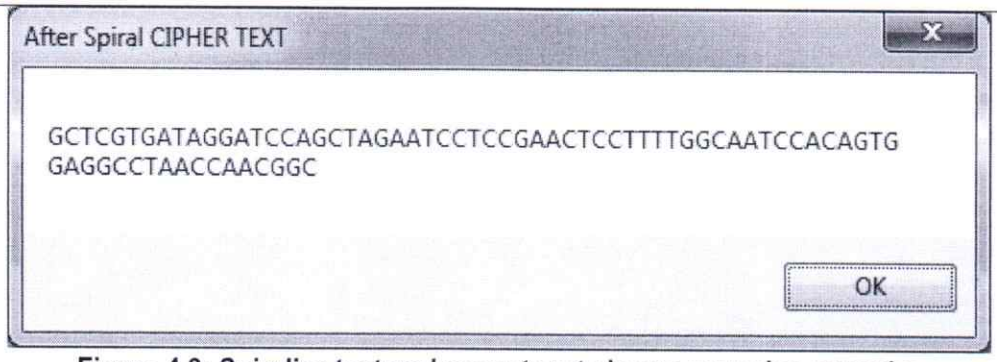



**Figure 4.9: Spiralize text and concatenate bases row wise row wise**

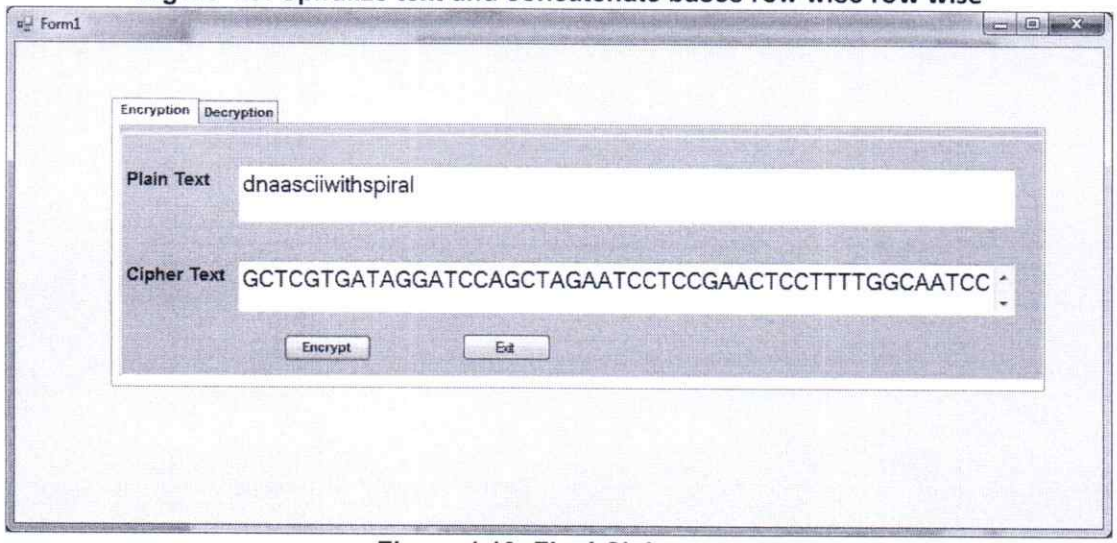



**Figure 4.10: Final Cipher text**

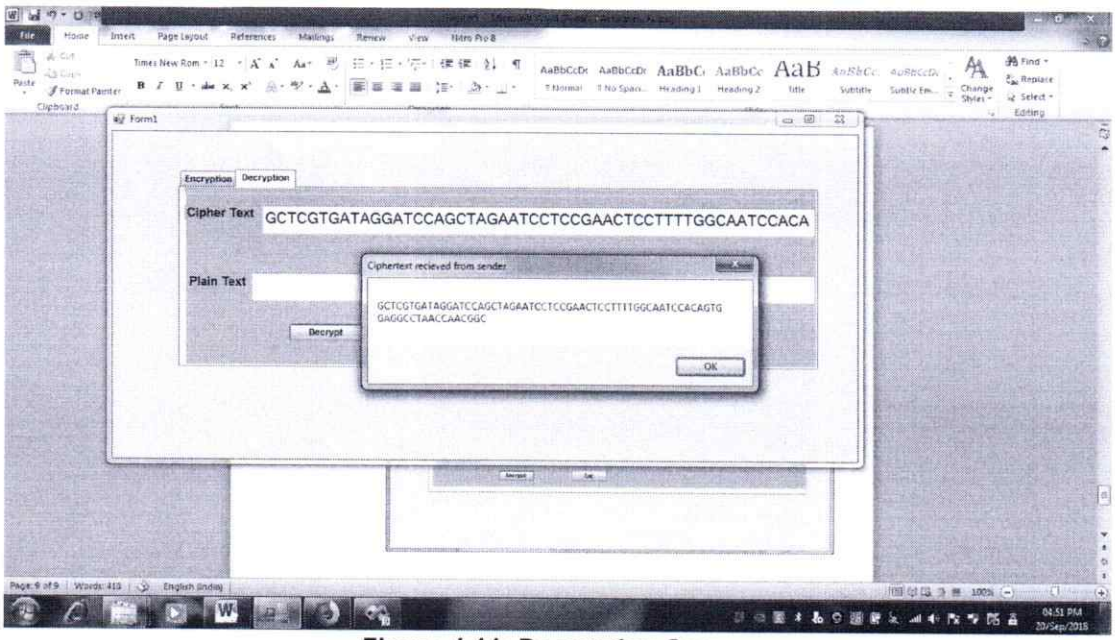



**Figure 4.11: Decryption Screen**

PRINCIPAL
MALINENI LAKSHMAIAH
WOMEN'S ENGINEERING COLLEGE
PULLADIGUNTA, GUNTUR-17.


DOI: 10050086.2022.11.24

ISSN: 1005-0086

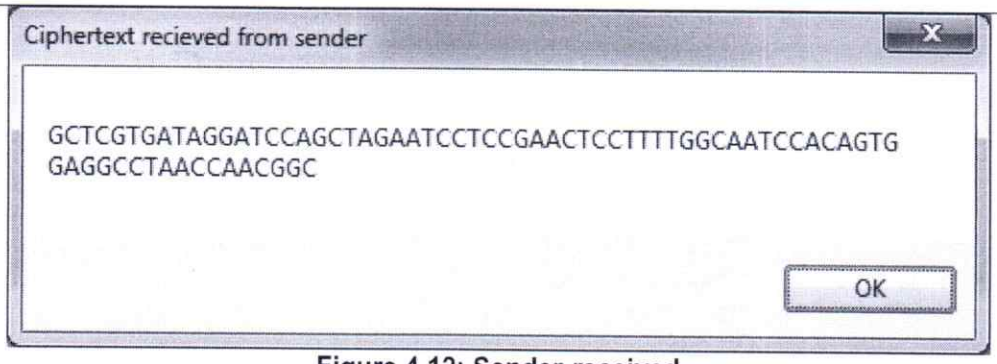



Figure 4.12: Sender-received

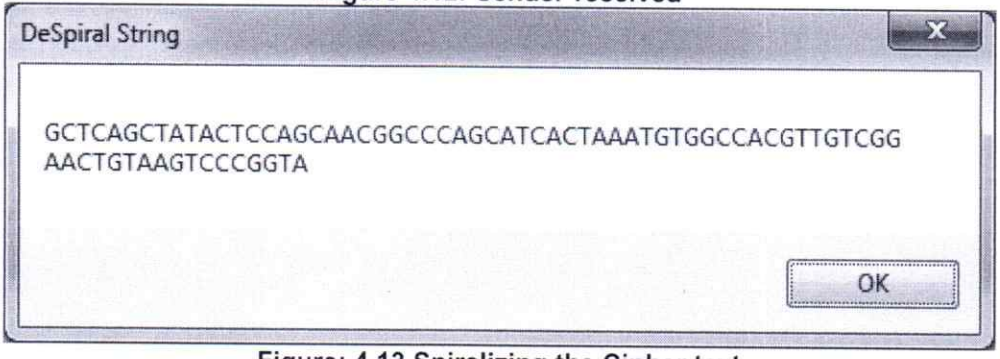



Figure: 4.13 Spiralizing the Cipher text

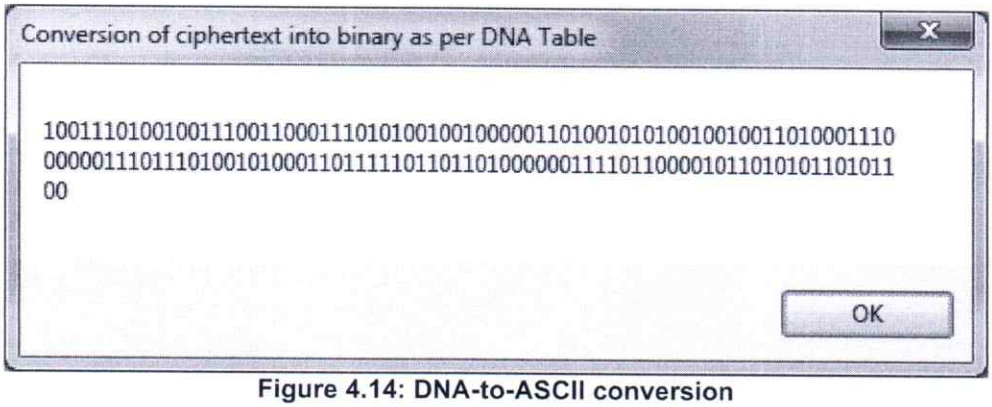



Figure 4.14: DNA-to-ASCII conversion

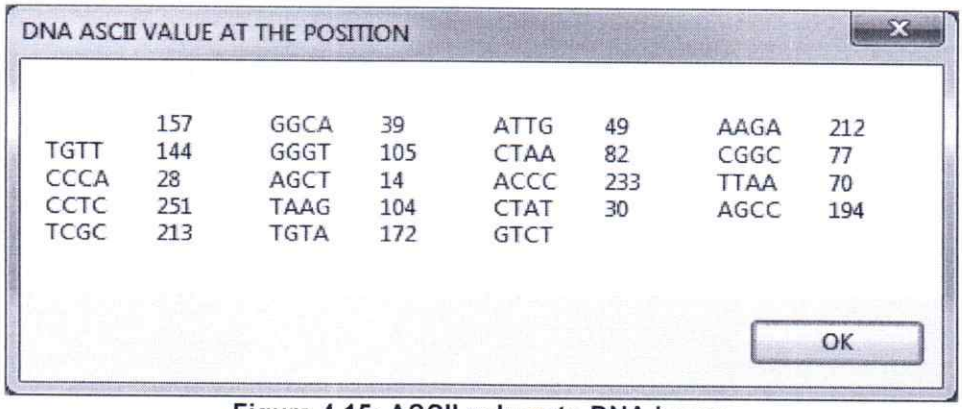



Figure 4.15: ASCII values to DNA bases

PRINCIPAL
MALINENI LAKSHMAIAH
WOMEN'S ENGINEERING COLLEGE
PULLADIGUNTA, GUNTUR-13


DOI: 10050086.2022.11.24

ISSN: 1005-0086

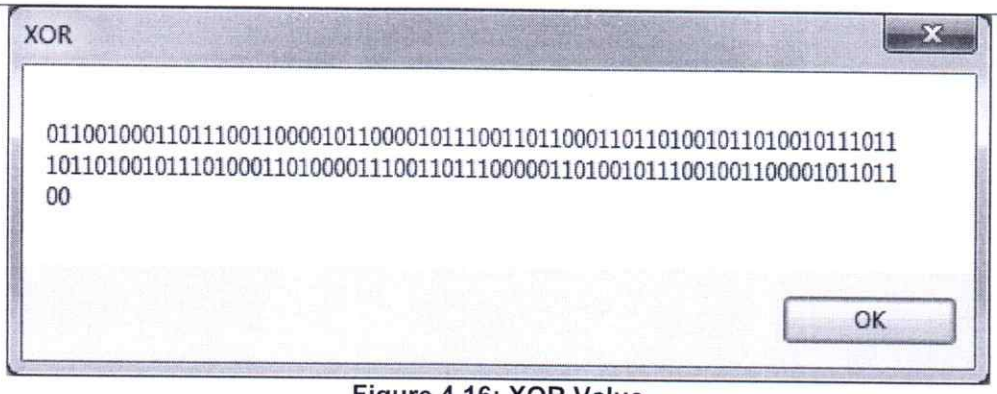



**Figure 4.16: XOR Value**

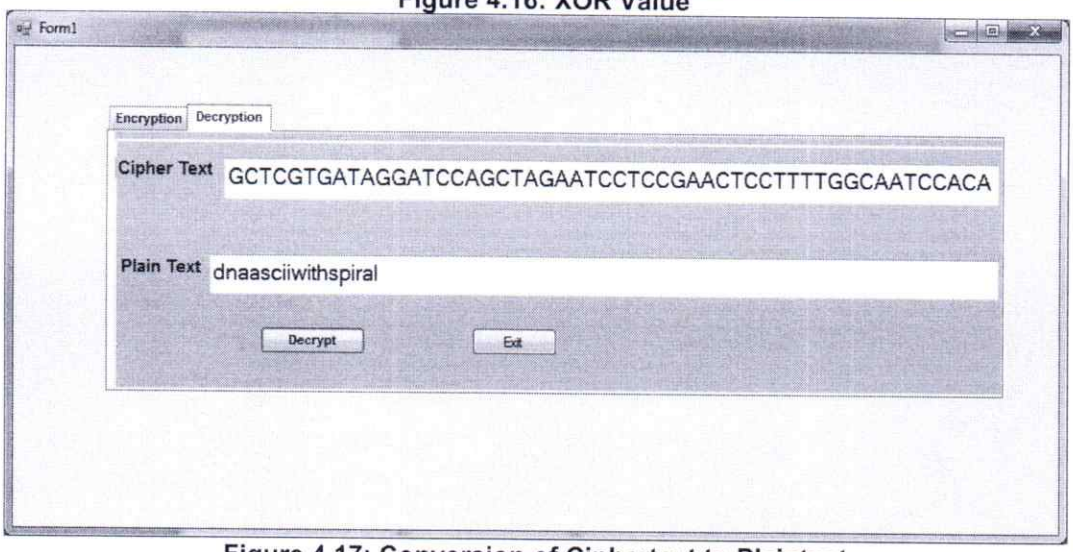



**Figure 4.17: Conversion of Ciphertext to Plaintext**

### a. Performance Measurement

Analyzing an algorithm measured on Intel i3 CPU in.NET environment for the text "dnaasciiwithspiral". The simulations recorded MB message lengths and encryption/decryption times. Compared to the prior suggested model "Level Based DNA Security Mechanism utilizing DNA Codons," this model fits all message lengths and executes quickly.

**Table 4.1.1: Encryption time comparison between proposed and level-based DNA Cryptosystems employing DNA Codons**

| Sl.No. | Message Length(in terms of MB) | Proposed Algorithm Encryption Time (ms). | Codon-based DNA cryptosystem encryption time (ms). |
|---|---|---|---|
| 1 | 10 | 0.001856 | 0.00421 |
| 2 | 100 | 0.00495 | 0.01234 |
| 3 | 1000 | 0.0301 | 0.123 |
| 4 | 10000 | 0.08765 | 13.064 |
| 5 | 20000 | 0.18768 | 22.28 |
| 6 | 30000 | 0.8567 | 30.334 |


DOI: 10050086.2022.11.24

ISSN: 1005-0086

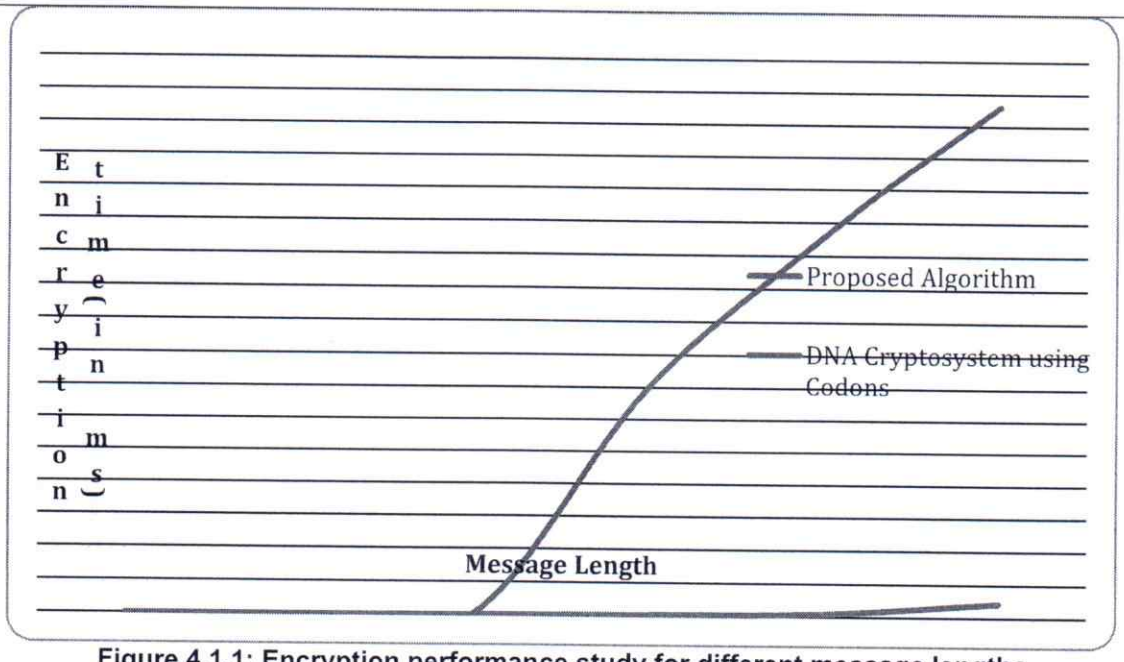



Figure 4.1.1: Encryption performance study for different message lengths

Table4.1.2: Analysis of encryption performance for different message lengths

| Sl.No. | Message Length(in terms of MB) | Proposed Algorithm Decryption Time (ms). | Codons-based DNA cryptosystem decryption time (ms). |
|---|---|---|---|
| 1 | 10 | 0.0001768 | 0.0001542 |
| 2 | 100 | 0.0010847 | 0.0005971 |
| 3 | 1000 | 0.0159 | 0.00191 |
| 4 | 10000 | 0.134 | 0.0299 |
| 5 | 20000 | 1.08 | 3.61 |
| 6 | 30000 | 4.64 | 8.123 |

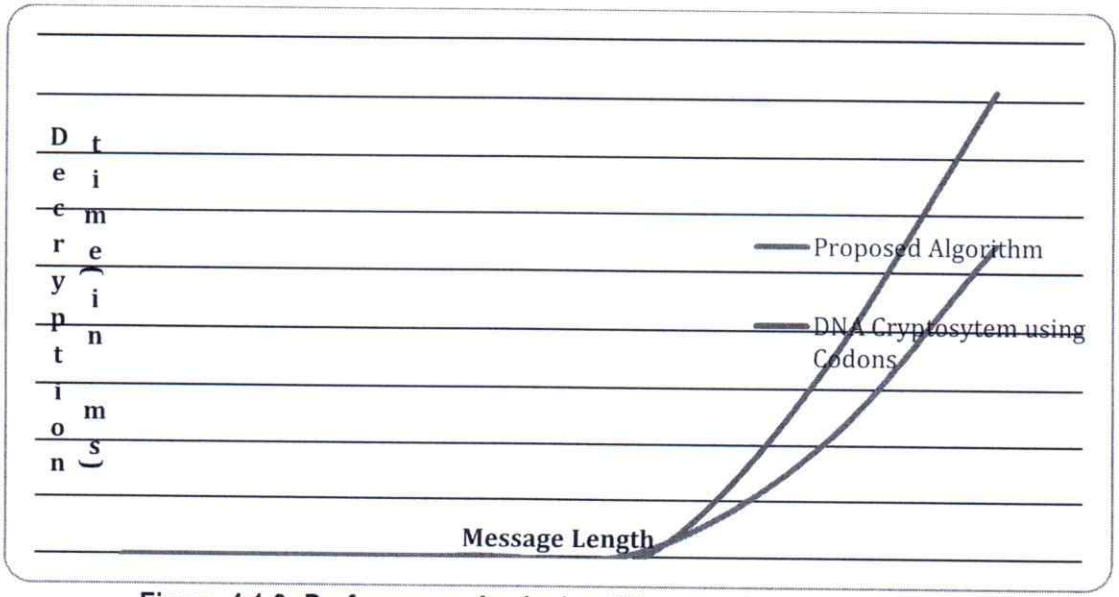



Figure 4.1.2: Performance Analysis of Message Length Decryption

## Avalanche Effect


DOI: 10050086.2022.11.24

PRINCIPAL
MALINENI LAKSHMAIAH
WOMEN'S ENGINEERING COLLEGE
PULLADIGUNTA, GUNTUR-17.

ISSN: 1005-0086

The following table represents the percentage of bits flipped in the cipher text when a single character in the plaintext is modified.
Plaintext1= national
Plaintext2=natioxal
Ciphertext1:
01101001100000010001010011000101000010111110110101100100011000 11
Ciphertext2:10110001101110111111111111010011000100101010111101011111110010110

**Table 4.2.1 Comparison of Avalanche Effect for strings national and natioxal with Proposed model and Traditional Models.**

| S. No. | Name of the Algorithm | Number of bits Flipped | % |
|---|---|---|---|
| 1 | Ceaser Cipher | 2 | 3.125 |
| 2 | Playfair | 4 | 4.69 |
| 3 | Vigener Cipher | 3 | 3.125 |
| 4 | Blowfish | 19 | 29.68 |
| 5 | DES | 32 | 51.56 |
| 6 | Feistel Inspired DNA based Cryptosystem | 34 | 53.12 |
| 7 | DNA Cryptosystem using DNA ASCII Table with Spiral Approach | 40 | 62.5 |

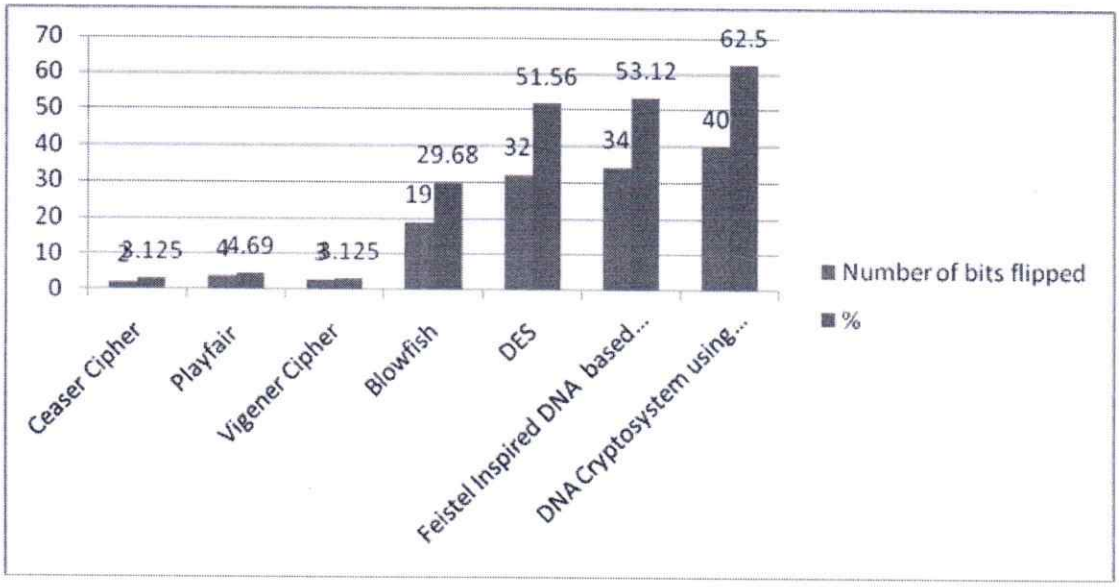



**Figure 4.2.1: Comparison of the results of the avalanche effect for strings national and natioxal**

In the considered plaintext "national", the change of one character leads to change in 40 bits in the ciphertext which was very difficult to the intruder in identifying the plaintext. It was observed that the total number of bits flipped in the present model is more when compared to previous proposed model, Inspired Feistel Cipher using D-Box.

## V. CONCLUSION

In the suggested approach, the intruder cannot figure out the plaintext since the key is generated randomly. The spiral design is first confirmed by transmitter and receiver. The sender's randomly generated key is sent to the recipient over the secure media. The method includes three levels of security. The primary level does the cross join of XOR with the plaintext and key, i.e., XOR(ML,KR) and XOR(MR,KL). The four-length DNA sequence is mapped to DNA ASCII value at the next level. Another DNA ASCII Table is intended to identify mapping in 256! ways, making intruder activities more complicated. The DNA sequences spiral in the final level. Row-wise information is provided to the recipient. To


DOI: 10050086.2022.11.24

increase security, the suggested method increased confusion and dispersion. This chapter also compares the Avalanche Effect on current encryption methods with the proposed technique.

## REFERENCES

1. Prasanna Balaji Narasingapuram, M. Ponnavaikko, "DNA Cryptography Based User Level Security for Cloud Computing and Applications," International Journal of Recent Technology and Engineering (IJRTE), ISSN: 2277-3878, Volume-8 Issue-5, January 2020.
2. Hamza Hammami, Hanen Brahmi, Sadok Ben Yahia,"Secured Outsourcing Towards a Cloud Computing Environment Based on DNA Cryptography," IEEE, pp. 31-36, 2018.
3. Bahubali Akiwate, Latha Parthiban, "Enhanced DNA Cryptographic Solution for Secured Data Transmission," International Journal of Innovative Technology and Exploring Engineering (IJITEE) ISSN: 2278-3075,2019.
4. Bahubali Akiwate, Latha Parthiban, "A Dynamic DNA for Key-based Cryptography," CTEMS, IEEE, 2018.
5. Kaur, Karandeep, "A Double Layer Encryption Algorithm based on DNA and RSA for Security on Cloud,"International Research Journal of Engineering and Technology, Volume: 03 Issue: 03 Mar-2016.
6. Himanshu Kumar Shukla, Satyam Dubey, "Open Access ArticleSecurity in Internet of Things (IoT) Hashing Cryptographic Functions", Vol.7, Issue.3, pp.7-11, Jun-2019.
7. S.V.Keerthana Priya, S.J.Saritha, "A Robust Technique to Generate Unique Code DNA Sequence," IEEE, pp. 3815-3820, 2017.
8. Zhang et al, "DNA based random key generation and management for OTP encryption," 2017 Sep;159:51-63. DOI: 10.1016/j.biosystems.2017.07.002. Epub 2017 Jul 18.
9. Hassan Al-Mahdi Meshrif Alruily Osama R.Shahin, Khalid Alkhaldi, "Design and Analysis of DNA Encryption and Decryption Technique based on Asymmetric Cryptography System," (IJACSA) International Journal of Advanced Computer Science and Applications, Vol. 10, No. 2, 2019.
10. M. Thangavel, P. Varalakshmi, "Enhanced DNA and ElGamal cryptosystem for secure data storage and retrieval in the cloud," Springer Science+Business Media, LLC, part of Springer Nature 2017. [11] Md. Rafiul Biswas, Kazi Md. Rokibul Alam, Ali Akber, and Yasuhiko Morimoto, "A DNA Cryptographic Technique Based on Dynamic DNA Encoding and Asymmetric Cryptosystem," IEEE, 978-1-5386-3288-8/17 ©2017.
11. A.Vyasa Bharadwaja, V. Ganesan, "DNA Computing Based Encryption Algorithm for Wireless Multimedia Communication System," International Journal of Innovative Technology and Exploring Engineering (IJITEE), ISSN: 2278-3075, Volume-9 Issue3, January 2020
Saifali Mavanai, Ajay Pal, Ravi Pandey, Asst Prof. Deepika Nadar, "Message Transmission Using DNA Crypto-System," International Journal of Computer Science and Mobile Computing.
12. Vinay S, Adarsh Pujar, Ankith, H.Akshay Kedlaya, Vasudev S Shahapur, " Implementation of DNA Cryptography based on Dynamic DNA Sequence Table using Cloud Computing", International Journal of Engineering Research & Technology (IJERT)ISSN: 2278-0181 Published by, RTESIT - 2019 Conference Proceedings
13. Gambhir Singh, Rakesh Kumar Yadav, "DNA Based Cryptography Techniques with Applications and Limitations,"International Journal of Engineering and Advanced Technology (IJEAT), ISSN: 2249 – 8958, Volume-8 Issue-6, August 2019.
14. Partha Sarathi Goswami, Tamal Chakraborty, Harekrishna Chatterjee, "A Novel Encryption Technique Using DNA Encoding and Single Qubit Rotations", Vol.6 , Issue.3 , pp.364-369, Mar-2018.
15. Santhi G., "Secure Data transmission using Recombinant DNA Cryptography and Morse Code Pattern", Vol.8 , Issue.5 , pp.174-177, May-2020.
16. Jignesh Patel, Foram Suthar, Samrat.V.O.Khanna, "Open Access ArticleA Critical Analysis on Encryption Techniques used for Data Security in Cloud Computing and IOT (Internet of Things) based Smart cloud storage System: A Survey", Vol.7, Issue.2, pp.21-25, Apr-2019.

PRINCIPAL
MALINENI LAKSHMAIAH
WOMEN'S ENGINEERING COLLEGE
PULLADIGUNTA, GUNTUR-17.


DOI: 10050086.2022.11.24

JOURNAL OF HARBIN INSTITUTE OF TECHNOLOGY
ISSN: 0367-6234 Vol. 54 Iss. 3 2022

# ADVANCED CYBER CRIME RATE PREDICTION USING DEEP LEARNING & AI BASED ALGORITHM

**Dr. M. Bheemalingaiah[*1], Dr. E. Nageswara Rao[2], Dr. Abburi Srirama Kanaka Ratnam[3] & Sudhakar Vecha[4]**

[*1]Professor, Dept. of CSE, Malineni Lakshmaiah Women's Engineering College, Guntur, A.P, India

[2&3]Associate Professor, Dept of CSE, Malineni Lakshmaiah Women's Engineering College, Guntur, A.P, India

[4]Assistant Professor and HOD, Dept. of CSE, Malineni Lakshmaiah Women's Engineering College, Guntur, A.P, India

**ABSTRACT**

Our Research "Advanced Cyber Crime Rate Prediction Using Deep Learning & AI Based Algorithm" is a wrongdoing is an intentional demonstration that can hurt, as well as property harm or misfortune, and can prompt discipline by a state or other authority as indicated by the seriousness of the wrongdoing. The number and types of crimes are expanding at a disturbing rate, constraining offices to foster proficient techniques to go to preventive lengths. In the current situation of quickly expanding wrongdoing, conventional wrongdoing tackling methods can't convey results, being slow paced and less productive. Consequently, in the event that we can think of ways of foreseeing wrongdoing, exhaustively, before it happens, or concoct a "machine" that can help cops, it would lift the weight of police and help in forestalling violations. To accomplish this, we propose including AI (ML) and PC vision calculations and procedures. In this paper, we depict the consequences of specific situations where such methodologies were utilized, and which spurred us to seek after additional examination in this field. The primary justification for the adjustment of wrongdoing identification and counteraction lies in the when measurable perceptions of the specialists utilizing such procedures. The sole motivation behind this study is to decide how a blend of ML and PC vision can be utilized by regulation organizations or specialists to recognize, forestall, and tackle wrongdoings at a substantially more exact and quicker rate. In synopsis, ML and PC vision strategies can achieve a development in regulation offices.

**KEYWORDS:** Advanced, Cyber, Crime, Rate, Prediction, Deep Learning, AI, Algorithm.

## 1. INTRODUCTION

Research is right now being led on the arrangement of numerical methods to recuperate and make it feasible for PCs to fathom 3D pictures. Acquiring the 3D visuals of an article assists us with object discovery, walker identification, face acknowledgment, Eigenfaces dynamic appearance and 3D shape models, individual photograph assortments, example acknowledgment, mathematical arrangement, enormous data sets, area acknowledgment, classification acknowledgment, pack of words, part-based models, acknowledgment with division, insightful photograph altering, setting and scene understanding, and huge picture assortment and learning, picture look, acknowledgment information bases, and test sets.

These are just fundamental applications, and every classification referenced above can be additionally investigated. In ref. [4], VLFeat is presented, which is a library of PC vision calculations that can be utilized to lead quick prototyping in PC vision research, along these lines empowering an instrument to acquire PC vision results a lot quicker than expected. Thinking about face discovery/human acknowledgment [5], human stance can likewise be perceived. In this way, PC vision is incredibly appealing for envisioning our general surroundings. Albeit different wrongdoings and their basic nature appear to be flighty, how unforeseeable would they say they are? In ref. [3], the creators brought up that as society and the economy brings about new sorts of violations, the requirement for an expectation framework has developed. In ref. [2], wrongdoing patterns and expectation innovation called Mahanalobis and a unique time wrapping strategy are given, conveying the chance of foreseeing wrongdoing and securing the real guilty party. As depicted in ref. [3], in 2020, the United States National Institute of Justice conceded five awards for wrongdoing guaging as an augmentation to wrongdoing planning.

DOI: 10.11720/JHIT.54032022.17

Utilizations of wrongdoing guaging are right now being involved by regulation authorization in the United States, the United Kingdom, the Netherlands, Germany, and Switzerland [4]. These days, criminal mind with the assistance of advances in innovation is improving as time passes. Thus, it has become vital for us to give the police division and the public authority with the method for a new and strong machine (a bunch of projects) that can help them in their course of settling violations.

## 2. TECHNOLOGY

Wrongdoing estimating alludes to the fundamental course of foreseeing violations before they happen. Instruments are expected to anticipate a wrongdoing before it happens. Right now, there are apparatuses involved by police to aid explicit undertakings, for example, tuning in on a presumes call or utilizing a body cam to record some uncommon criminal behavior. Beneath we show whatever apparatuses to all the more likely comprehend where they could remain with extra mechanical help.

One great approach to following telephones is using a stingray [5], which is another outskirts in police observation and can be utilized to pinpoint a cell phone area by impersonating cell phone pinnacles and broadcasting the signs to deceive cellphones inside the area to communicate their area and other data. A contention against the utilization of stingrays in the United States is that it abuses the fourth amendment. This innovation is utilized in 23 states and in the area of Columbia.

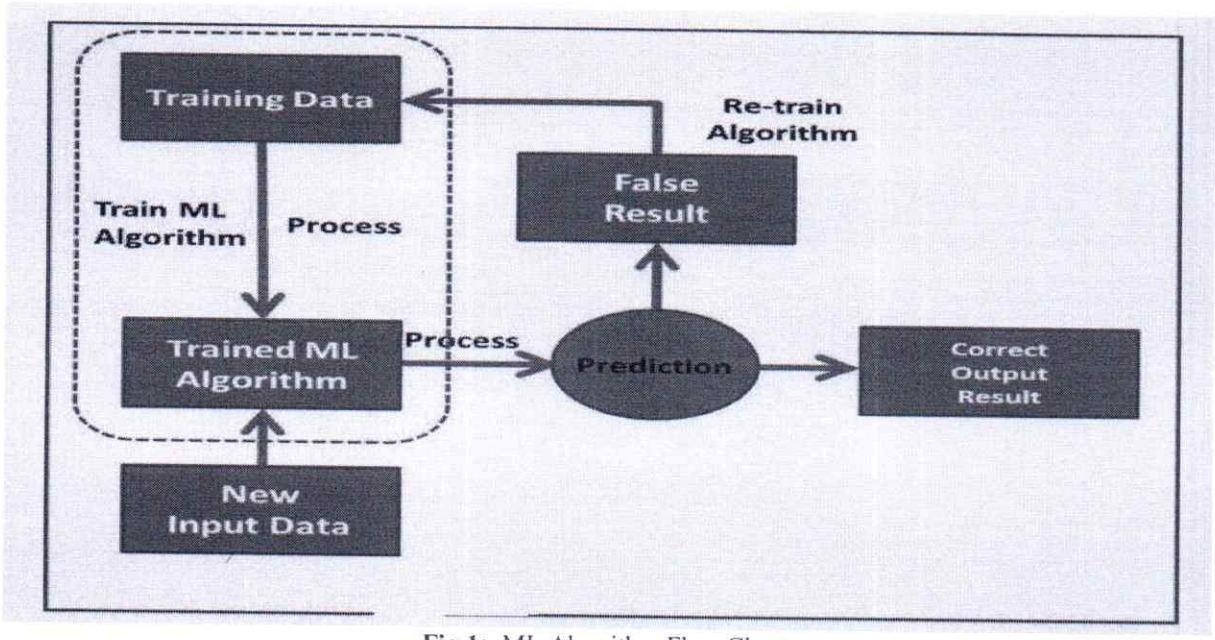



**Fig.1:** ML Algorithm Flow Chart

In ref. [6], the creators give knowledge on how this is something other than an observation framework, raising worries about security infringement. Furthermore, the Federal Communications Commission became involved and eventually encouraged the producer to meet two circumstances in return for an award: (1) "The promoting and offer of these gadgets will be restricted to government, state, nearby open security and regulation requirement authorities as it were" and (2) "State and neighbourhood regulation requirement organizations should propel coordinate with the FBI the obtaining and utilization of the gear approved under this approval." Although its utilization is advantageous, its execution remains incredibly disputable.

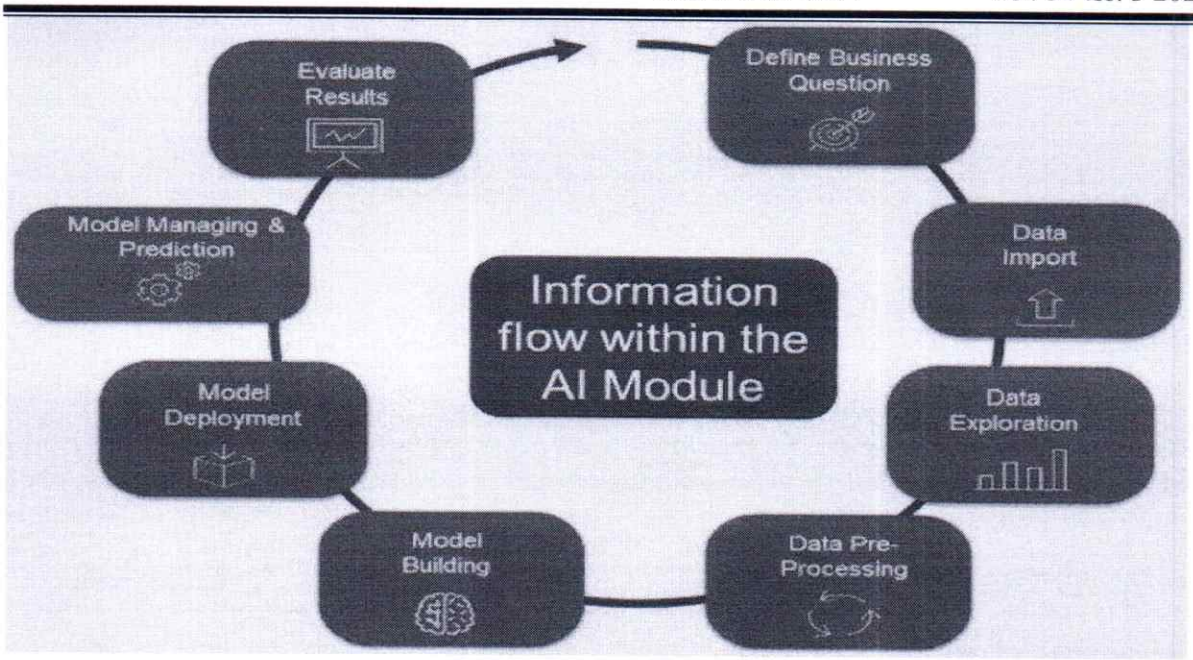



**Fig.2**: AI Based m, Flow chart.

Other observation techniques incorporate face acknowledgment, tag acknowledgment, and body cams. In ref. [3][2], the creators showed that facial acknowledgment can be utilized to acquire the profile of suspects and investigate it from various data sets to get more data. Likewise, a tag peruser can be utilized to get to information about a vehicle perhaps associated with a wrongdoing. They might even utilize body cams to see beyond what the natural eye can see, implying that the peruser notices all that a cop sees and records it. Regularly, when we see an article, we can't recall its total picture.

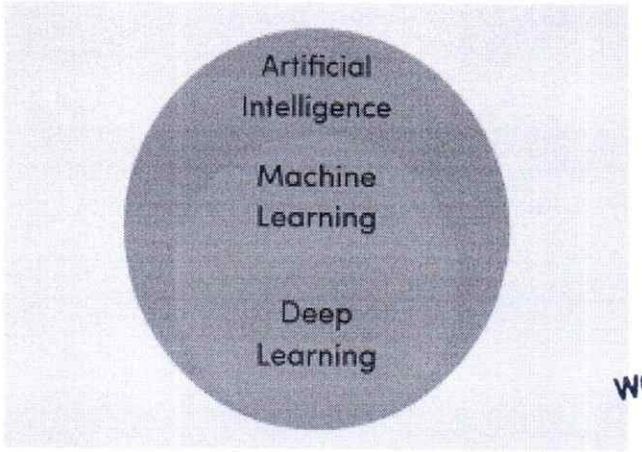



**Fig.3:** Deep Learning & AI Based.



In ref. [4][6], the effect of body cams was examined as far as official offense and abusive behavior at home when the police are making a capture. Body cams are in this way being worn by watch officials. In ref. [4][2]], the creators additionally referenced how assurance against unjust police rehearses is given. Notwithstanding, the utilization of body cams doesn't stop here, as other essential explanations behind having a body camera on consistently is to keep the happenings before the wearer with at least some expectations of record helpful occasions during every day exercises or during significant activities.

## 3. DL/ML- TECHNIQUE

The Communities and Crime Normalized Dataset versus genuine wrongdoing measurable information involving the open source information digging programming Waikato Environment for Knowledge Analysis (WEKA). Three calculations, specifically, straight relapse, added substance relapse, and choice stump, were carried out utilizing similar limited set of elements on networks and genuine wrongdoing datasets.

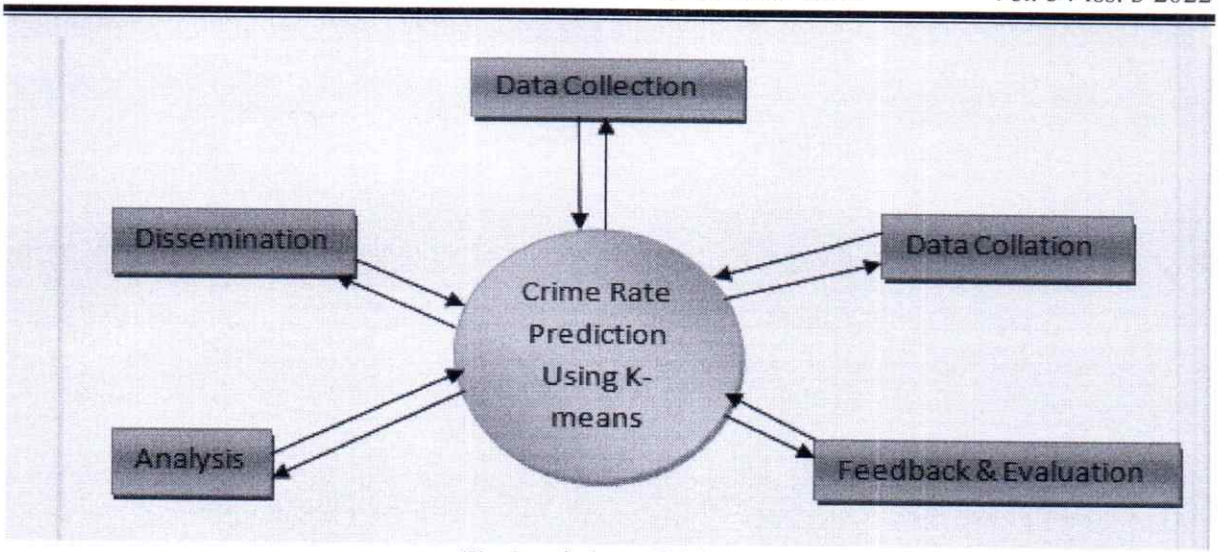



**Fig.4**: techniques K-Means

Test tests were arbitrarily chosen. The direct relapse calculation could deal with arbitrariness partially in the test tests and in this way ended up being awesome among each of the three chose calculations. The extent of the undertaking was to demonstrate the effectiveness and exactness of ML calculations in foreseeing rough wrongdoing designs and different applications, for example, deciding criminal areas of interest, making criminal profiles, and learning criminal patterns.

While considering WEKA [4][1][3]], the reconciliation of another graphical connection point called Knowledge Flow is conceivable, which can be utilized as a substitute for Internet Explorer.

IT gives a more focused perspective on information mining in relationship with the cycle direction, in which individual learning parts (addressed by java beans) are utilized graphically to show a specific progression of data. The creators then portray another graphical connection point called an experimenter, which as the name recommends, is intended to analyze the presentation of various learning plans on numerous informational indexes. Wrongdoing expectations were explored in light of ML. Wrongdoing information of the last 15 years in Vancouver (Canada) were investigated for expectation. This AI based wrongdoing examination includes the assortment of information, information order, distinguishing proof of examples, forecast, and perception. K-closest neighbor (KNN) and helped choice tree calculations were additionally carried out to examine the wrongdoing dataset.

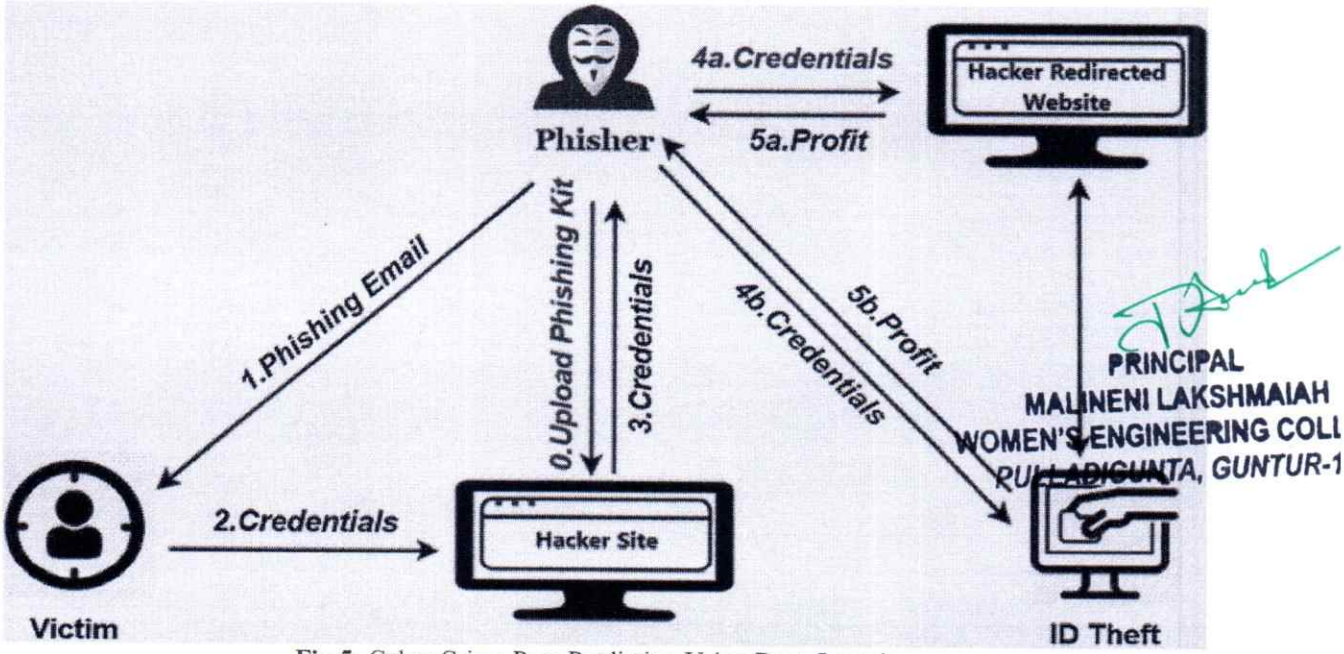



**Fig.5:** Cyber Crime Rate Prediction Using Deep Learning

In their review, a sum of 560,000 wrongdoing datasets somewhere in the range of 2003 and 2018 were broke down, and wrongdoing expectation with an exactness of somewhere in the range of 40% and 44% was acquired by anticipating the wrongdoing utilizing ML calculations. The exactness was low as a forecast model, yet the creators inferred that the precision can be expanded or improved by tuning both the calculations and wrongdoing information for explicit applications.

ML a

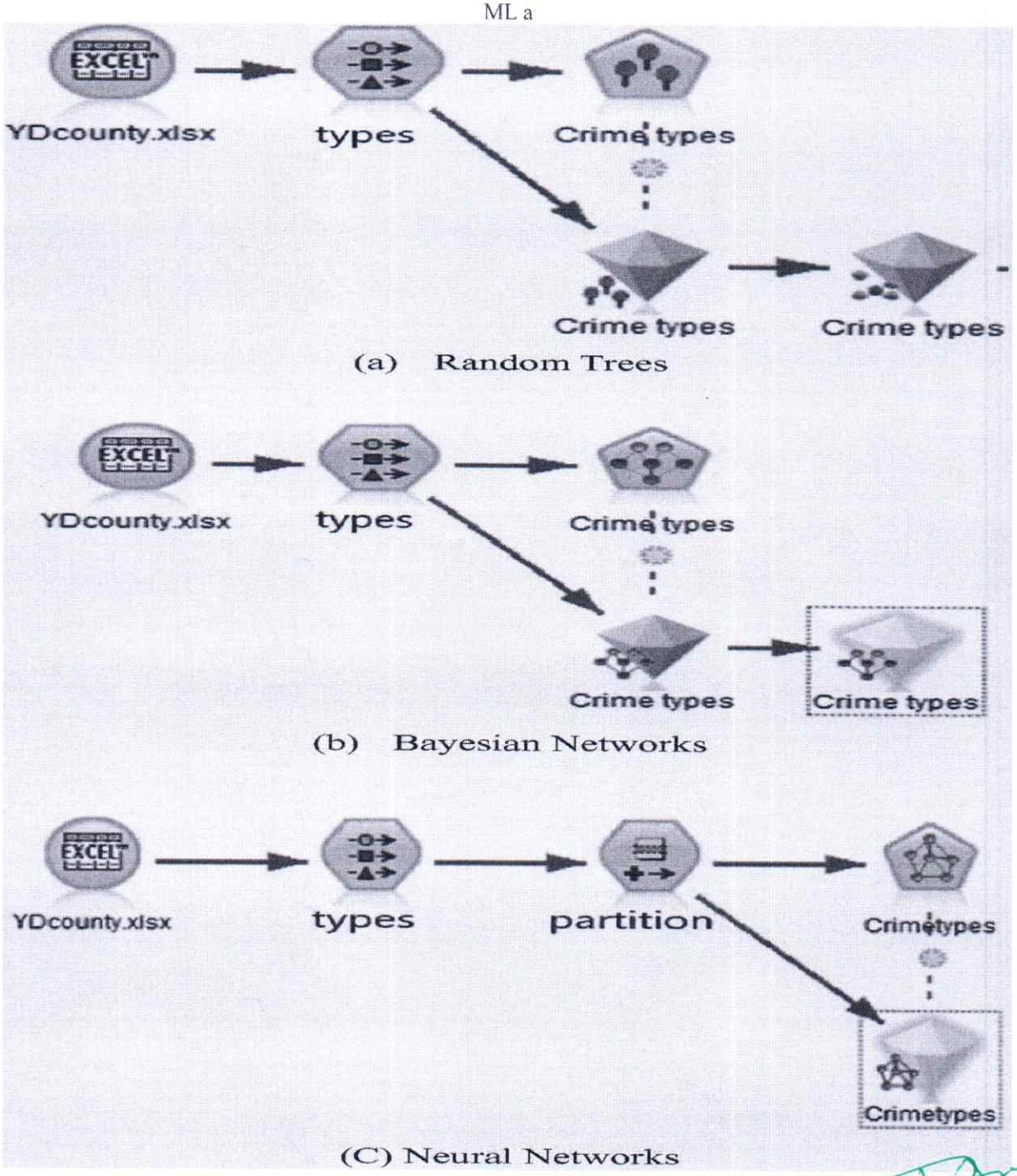



**Fig.6:** NN

Ind information science strategies were utilized for wrongdoing forecast in a wrongdoing dataset from Chicago, United States. The wrongdoing dataset comprises of data like the wrongdoing area portrayal, sort of wrongdoing, date, time, and exact area facilitates. Various mixes of models, for example, KNN grouping, strategic relapse, choice trees, irregular woodland, a help vector machine (SVM), and Bayesian techniques were tried, and the most dependable model was utilized for preparing. The KNN arrangement ended up being awesome with a precision of around 0.7857.

They additionally utilized various charts that aided in understanding the different qualities of the wrongdoing dataset of Chicago. The principle motivation behind this paper is to give a thought of how ML can be utilized by regulation requirement offices to anticipate, distinguish, and settle wrongdoing at a greatly improved rate, which brings about a decrease in wrongdoing.

an element level information combination strategy in light of a profound brain organization (DNN) is proposed to precisely anticipate wrongdoing event by effectively melding multi-model information from a few areas with ecological setting data. The dataset comprises of information from a web-based data set of wrongdoing measurements from Chicago, segment and meteorological information, and pictures. Wrongdoing expectation strategies use a few ML methods, including a relapse examination, portion thickness assessment (KDE), and SVM.

Their methodology essentially comprised of three stages: assortment of information, examination of the connection between wrongdoing episodes and gathered information utilizing a measurable methodology, and in conclusion, precise expectation of wrongdoing events. The DNN model comprises of spatial highlights, transient elements, and natural setting. The SVM and KDE models had exactness's of 67.041% and 66.353%, separately, while the proposed DNN model had a shocking precision of 84.235%. The test results showed that the proposed DNN model was more exact in foreseeing wrongdoing events than the other forecast models.

## 4. PROPOSED IDEA

Subsequent to finding and understanding different unmistakable techniques involved by the police for reconnaissance purposes, we decided the significance of every strategy. Every reconnaissance technique can perform well all alone and produce agreeable outcomes, despite the fact that for only one explicit trademark, or at least, assuming we utilize a Sting Ray, it can help us just when the suspect is utilizing a telephone, which ought to be turned on.

Hence, it is just valuable when the data with respect to the stake out area is right. In light of this data, we can perceive how the always developing innovation has once more created a shrewd method for leading observation. The presentation of profound learning, ML, and PC vision methods has given us another viewpoint on ways of directing observation. This is a canny way to deal with observation since it attempts to imitate a human methodology, yet it does so 24 h per day, 365 days every year, and whenever it has been shown how to do things it destroys them a similar way over and again.

## 5. CONCLUSIONS

Anticipating wrongdoings before they happen is easy to comprehend, however it takes significantly more than understanding the idea to make it a reality. This paper was composed to help analysts planning to make wrongdoing expectation a reality and carry out such trend setting innovation, all things considered. In spite of the fact that police in all actuality do incorporate the utilization of new innovations, for example, Sting Rays and facial acknowledgment like clockwork, the execution of such programming can essentially impact the manner in which police work, in a vastly improved way.

This paper illustrated a structure conceiving how the parts of machine and profound learning, alongside PC vision, can assist with making a framework that is significantly more accommodating to the police. Our proposed framework has an assortment of innovations that will perform everything from checking wrongdoing areas of interest to perceiving individuals from their voice notes. The main trouble confronted will be to really make this framework, trailed by issues like its execution and use, among others. Nonetheless, these issues are reasonable, and we can likewise profit from a security framework that screens the whole city nonstop. All in all, to imagine a reality where we join such a framework into a police power, tips or leads considerably more solid can be accomplished and maybe wrongdoing can be annihilated at a lot quicker rate.



## REFERENCES

1. D, Dixit R, Shah A, Shah P, Shah M (2020) A comprehensive analysis regarding several breakthroughs based on computer intelligence targeting various syndromes. Augment Hum Res 5(1):1. https://doi.org/10.1007/s41133-020-00033-z
2. Patel H, Prajapati D, Mahida D, Shah M (2020) Transforming petroleum downstream sector through big data: a holistic review. J Pet Explor Prod Technol 10(6):2601–2611. https://doi.org/10.1007/s13202-020-00889-2
3. Szeliski R (2010) Computer vision: algorithms and applications. Springer-Verlag, Berlin, pp 1–979

4. Vedaldi A, Fulkerson B (2010) Vlfeat: an open and portable library of computer vision algorithms. Paper presented at the 18th ACM international conference on multimedia. ACM, Firenze. https://doi.org/10.1145/1873951.1874249
5. Le TL, Nguyen MQ, Nguyen TTM (2013) Human posture recognition using human skeleton provided by Kinect. In: Paper presented at the 2013 international conference on computing, management and telecommunications. IEEE, Ho Chi Minh City. https://doi.org/10.1109/ComManTel.2013.6482417
6. Ahir K, Govani K, Gajera R, Shah M (2020) Application on virtual reality for enhanced education learning, military training and sports. Augment Hum Res 5(1):7. (https://doi.org/10.1007/s41133-019-0025-2)
7. Talaviya T, Shah D, Patel N, Yagnik H, Shah M (2020) Implementation of artificial intelligence in agriculture for optimisation of irrigation and application of pesticides and herbicides. Artif Intell Agric 4:58–73. https://doi.org/10.1016/j.aiia.2020.04.002
8. Jha K, Doshi A, Patel P, Shah M (2019) A comprehensive review on automation in agriculture using artificial intelligence. Artif Intell Agric 2:1–12. https://doi.org/10.1016/j.aiia.2019.05.004
9. Kakkad V, Patel M, Shah M (2019) Biometric authentication and image encryption for image security in cloud framework. Multiscale Multidiscip Model Exp Des 2(4):233–248. https://doi.org/10.1007/s41939-019-00049-y
10. Pathan M, Patel N, Yagnik H, Shah M (2020) Artificial cognition for applications in smart agriculture: a comprehensive review. Artif Intell Agric 4:81–95. https://doi.org/10.1016/j.aiia.2020.06.001.

JOURNAL OF HARBIN INSTITUTE OF TECHNOLOGY
ISSN: 0367-6234 Vol. 54 Iss. 3 2022

# ADVANCED FACE DETECTION USING MACHINE LEARNING &AI BASED ALGORITHM.

**Dr. M. Bheemalingaiah*1, Dr. Abburi Srirama Kanaka Ratnam[2], Venkaiah Chowdary. Bhimineni[3] & Dr. G. Ramaswamy[4]**

[*1]Professor, Dept. of CSE, Malineni Lakshmaiah Women's Engineering College, Guntur, A.P, India

[2&3]Associate Professor, Dept of CSE, Malineni Lakshmaiah Women's Engineering College, Guntur, A.P, India

[4]Professor and HOD, Dept. of CSE, Malineni Lakshmaiah Women's Engineering College, Guntur, A.P, India

**ABSTRACT**

Our Research "Advanced Face Detection Using Machine Learning &AI Based Algorithm" is a facial acknowledgment framework utilizing AI, explicitly support vector machines (SVM). The initial step required is face recognition which we achieve utilizing a broadly utilized strategy called the Viola-Jones calculation. The Viola-Jones calculation is exceptionally attractive because of its high location rate and quick handling time. When the face is identified, include extraction on the face is performed utilizing histogram of arranged slopes (HOG) which basically stores the edges of the face as well as the directionality of those edges. Hoard is a viable type of element extraction due its superior execution in normalizing neighbourhood contrast. Finally, preparing and grouping of the facial information bases is finished utilizing the multi-class SVM where every exceptional face in the facial data set is a class. We endeavor to utilize this facial acknowledgment framework on two arrangements of information bases, the AT&T face data set and the YALE B face data set and will examine the outcomes.

**KEYWORDS:** Advanced, Face, Detection, Machine, Learning, AI, Algorithm.

## 1. INTRODUCTION

Access control incorporates workplaces, PCs, telephones, ATMs, and so on the majority of these structures right now don't involve face acknowledgment as the standard type of giving section, yet with propelling advances in PCs alongside more refined calculations, facial acknowledgment is acquiring a few foothold in supplanting passwords and unique mark scanners. Since the time the occasions of 9/11 there has been a more concerned accentuation on creating security frameworks to guarantee the wellbeing of blameless residents. Specifically in spots, for example, air terminals and line intersections where distinguishing proof confirmation is vital, face acknowledgment frameworks conceivably can relieve the gamble and eventually keep future assaults from happening.

With respect to reconnaissance frameworks, a similar point can Peter Neal Barrina UCSD pbarrina@ucsd.edu be made assuming there are crooks running wild. Observation cameras with face acknowledgment capacities can helper in endeavors of finding these people. Then again, these equivalent observation frameworks can likewise assist with recognizing the whereabouts of missing people, albeit this is subject to powerful facial acknowledgment calculations as well as a completely evolved information base of appearances.

Also, ultimately, facial acknowledgment has surfaced in web-based entertainment applications on stages, for example, Facebook which propose clients to label companions who have been distinguished in pictures. Obviously there are numerous applications the purposes for facial acknowledgment frameworks. In everyday the means to accomplish this are the accompanying: face location, highlight extraction, and finally preparing a model.



DOI:10.[illegible]

## 2. FACE DETECTION

Facial location through the Viola-Jones calculation is a common method used because of its high identification rate and quick handling speed. The calculation can be summarized in four stages: include determination, highlight assessment, include figuring out how to make a classifier, and falling classifiers. Straightforward highlights are utilized, motivated by Haar premise capacities, which are basically rectangular elements in different arrangements.

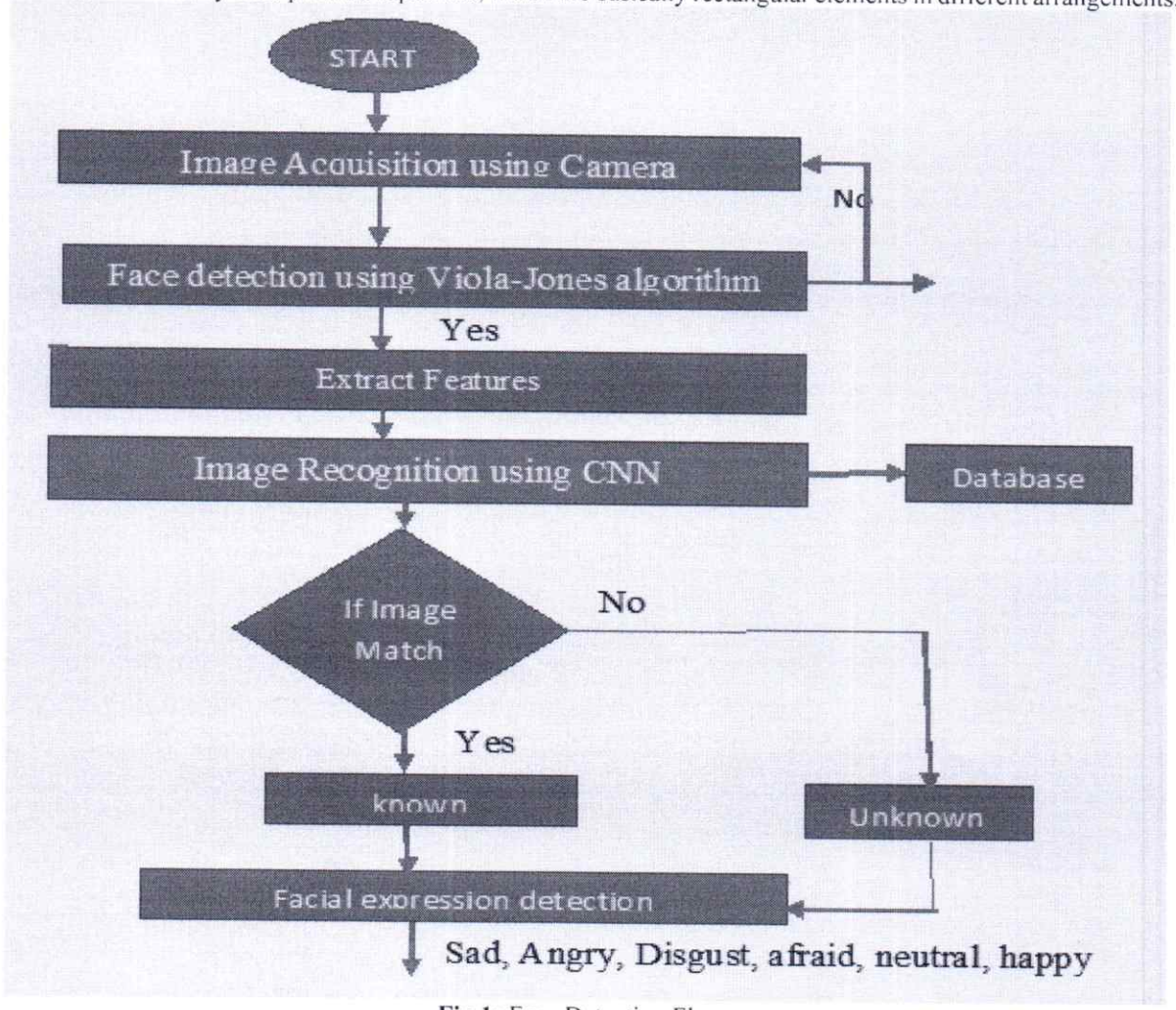



**Fig.1:** Face Detection Flow.

A two-square shape include addresses the distinction between the amount of the pixels in two adjoining districts of indistinguishable shape and size. This thought can be stretched out to the three-square shape and four-square shape highlights. To rapidly figure these square shape includes, a substitute portrayal of the information picture is required, called a basic picture.

The learning piece of the face identification calculation utilizes AdaBoost which fundamentally utilizes a direct mix of feeble grouping capacities to make a solid classifier. Every arrangement not entirely settled by the perceptron which delivers the most reduced mistake. Notwithstanding, this is described as a feeble student since the characterization work doesn't arrange the information well.

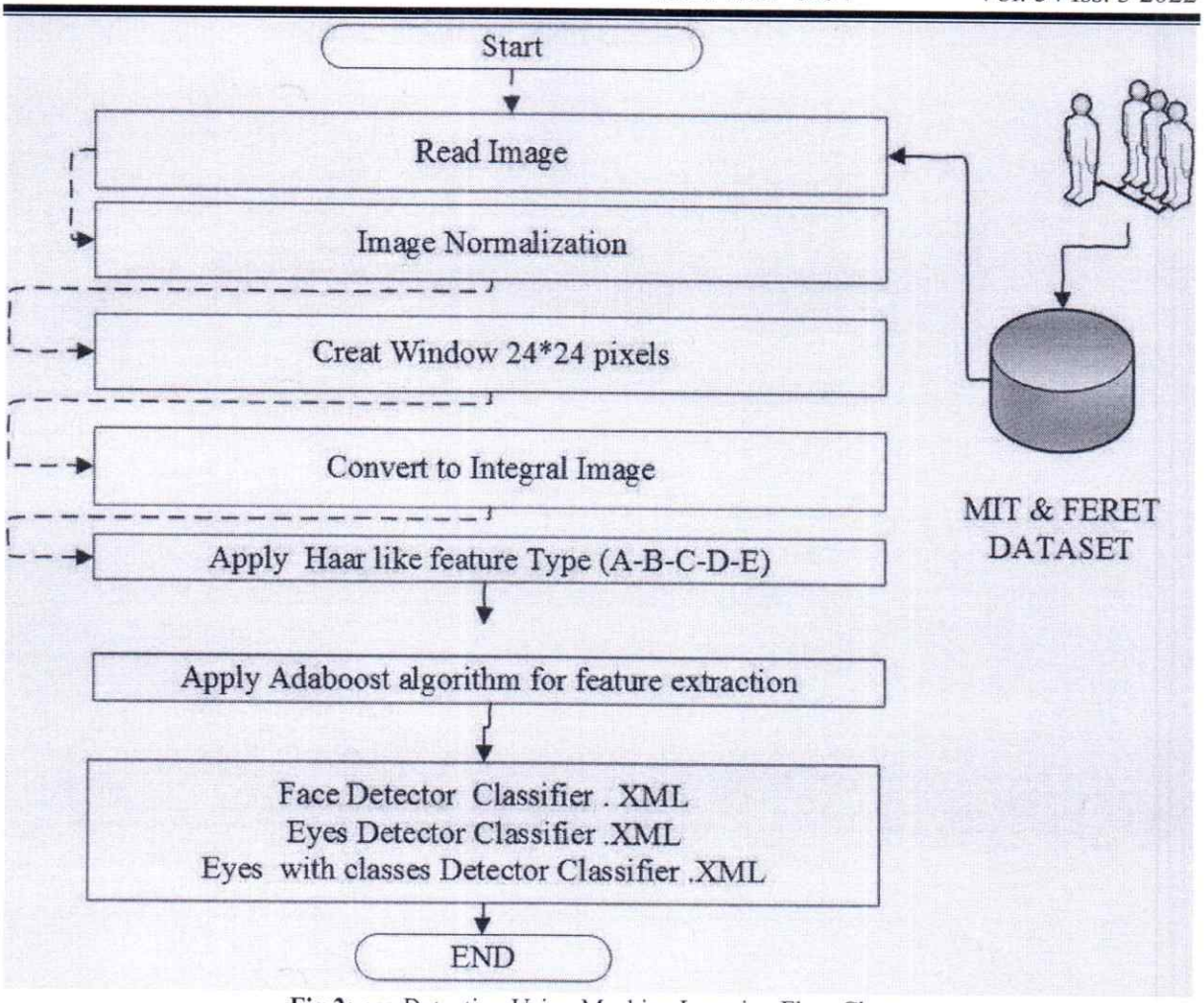



**Fig.2:** ace Detection Using Machine Learning Flow Chart.

To further develop results, a solid classifier is made after different rounds of re-weighting a set frail order capacities. These loads of the feeble arrangement capacities are conversely relative to their mistakes. The objective of this stage is to prepare the most applicable highlights of the face and to dismiss repetitive elements. The last advance of the Viola-Jones calculation is a course of classifiers. The classifiers developed in the past advance structure a course. Here up structure, the objective is to limit the calculation time and accomplish high discovery rate. Sub-windows of the info picture will be resolved a face or non-face with classifiers of expanding intricacy.

## 3. FEATURE EXTRACTION

Before we can recognize faces, it is first important to determine what elements of the face ought to be utilized to prepare a model. When the Viola-Jones face identification runs, the face part of the picture is then utilized for include extraction. It is critical to choose highlights which are novel to each face which are then used to store discriminant data in minimized include vectors. These component vectors are the critical piece of the preparation part of the facial acknowledgment framework and in our work we propose utilizing HOG highlights.

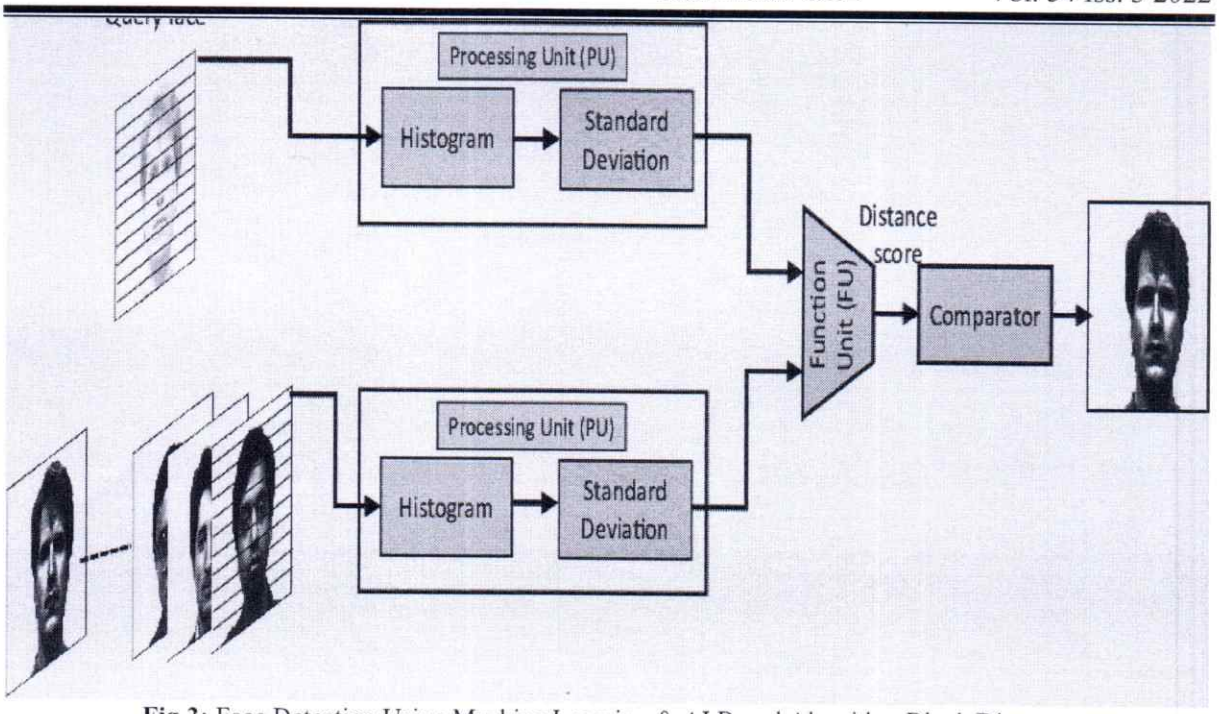



**Fig.3:** Face Detection Using Machine Learning & AI Based Algorithm Block Diagram

As referenced already, HOG highlights perform well since they store edges and edge course. Excellent neighbourhood contrast standardization, course spatial binning and fine direction binning are largely fundamental to great HOG results. Removing HOG elements can be summed up with the accompanying advances: compute angle of the picture, work out the histogram of slopes, standardize histograms, lastly structure the HOG include vector.

**Training Model**

When the technique for include extraction for a face has been laid out, the subsequent stage is to prepare a model involving the extricated highlight vectors of currently distinguished faces in an information base. The technique for preparing utilized endeavors to use different examples of each unmistakable face in the exhibition, to such an extent that the subsequent model will actually want to best match an information face to a personality from the display.

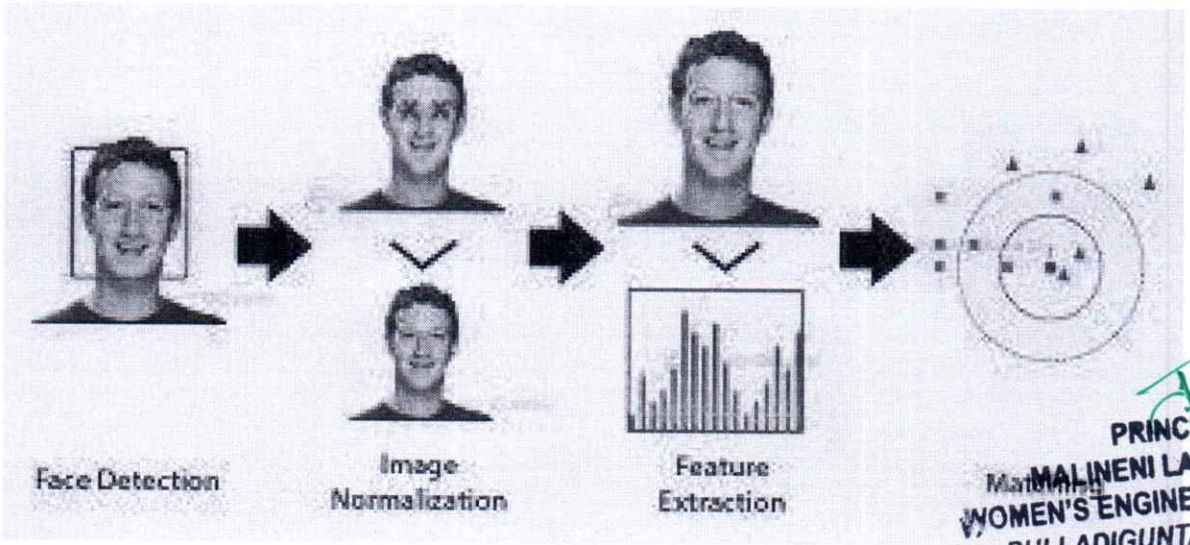





**Fig.4:** Support Vector Machines (SVM) Training.

Support Vector Machines (SVM) are a well-known preparing apparatus which can be utilized to create a model in light of a few classes of information, and afterward recognize them. For the fundamental two-class characterization issue, the objective of a SVM is to isolate the two classes by a capacity initiated from accessible

models. On account of facial acknowledgment, a class addresses an extraordinary face, and the SVM endeavors to find what best isolates the numerous component vectors of one interesting face from those of another exceptional face.

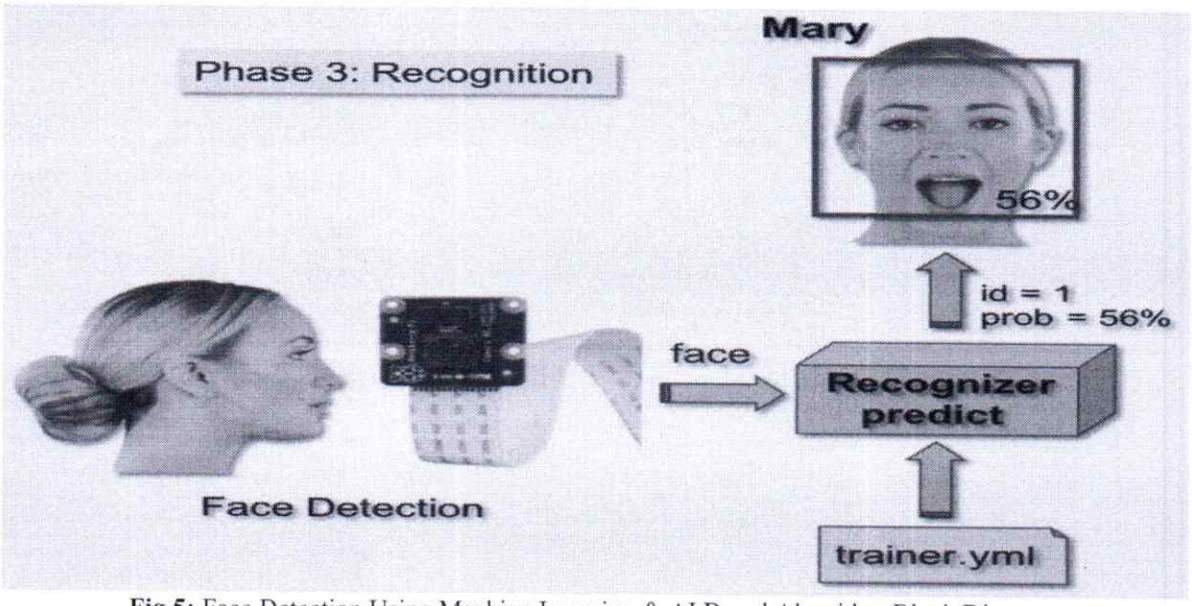



**Fig.5:** Face Detection Using Machine Learning & AI Based Algorithm Block Diagram.

Assuming we address all the element vectors of the two classes as elements with a similar dimensionality: (x1,y1),...,(xl, yl) where xi ∈ Rn, yi ∈ −1,1, and l is the all-out number of items, the SVM finds the hyperplane wx +b = 0 that isolates the biggest conceivable part of points of a similar class on the same side, while amplifying the separation from one or the other class to the hyperplane.

Since we have characterized the interaction for two-class acknowledgment utilizing SVM, this fundamental structure square can be stretched out to execute multi-class face acknowledgment. Expecting there are numerous classes in the informational index, a base up parallel tree is planned which applies a one-against-one methodology to order between various pairings of classes. For the model displayed in Figure 8, there are 8 classes which should be assessed given an info test face. Each part of the tree addresses a two-class SVM from which results one "victor" that go on up the tree.

**Dataset**

The first dataset we used to prepare and test the facial acknowledgment calculation was the AT&T dataset of countenances. The AT&T set is made out of forty special faces each with ten unique pictures with. The sythesis of the pictures is generally front facing view, predictable lighting conditions, and a few differing articulations. While preparing the model utilizing SVM, eight of the ten pictures for every individual were utilized to make a class for each face while the excess two pictures were held to test the model and the eighty pictures that were tried 73/80 countenances were effectively perceived bringing about a precision of 95.2545% which is alluring.

**Fig.6:** Date Set

**Yale B Dataset**

The second dataset we investigated testing was the Yale B Dataset. The subset we tried on included ten remarkable faces each with twenty pictures. The pictures are additionally front facing view without any progressions in articulation yet have changes in lighting force. While preparing the model utilizing SVM, sixteen of the twenty pictures were utilized to make a class for every individual while the excess four pictures were utilized for testing. Of the forty pictures which were tried, 40/41 countenances were effectively perceived which brings about a precision of 97.523%, an improvement from the past dataset. This improvement in exactness is because of two elements.

The principal motivation behind why the precision expanded is on the grounds that the pictures of each face were steady ready and looks not at all like the AT&T dataset which has shifting appearances. Furthermore, every special face has more pictures utilized in preparing along these lines the model is stronger and holds more discriminant data of each face in the exhibition.

The other methodology in the interim, which is part based, endeavors to manage present changes by permitting an adaptable mathematical connection between the parts in the order stage. For instance, by autonomously matching formats of the eyes, nose, and mouth rather than the whole face all in all, the design of the parts during order is unconstrained and doesn't uphold a particular mathematical model of the face. As opposed to the worldwide methodology where a face finder removes the face and engenders it to a bunch of SVMclassifiers, in the part approach, the face locator rather separates nearby parts of the face. Then, these neighbourhood parts are taken care of freely into a bunch of SVMclassifiers.

## 4. RESULTS

```
# Creating database
# It captures images and stores them in datasets
# folder under the folder name of sub_data
import cv2, sys, numpy, os
haar_file = 'haarcascade_frontalface_default.xml'

# All the faces data will be
#  present this folder
datasets = 'datasets'


# These are sub data sets of folder,
# for my faces I've used my name you can
```

```
# change the label here
sub_data = 'vivek'

path = os.path.join(datasets, sub_data)
if not os.path.isdir(path):
    os.mkdir(path)

# defining the size of images
(width, height) = (130, 100)

#'0' is used for my webcam,
# if you've any other camera
#  attached use '1' like this
face_cascade = cv2.CascadeClassifier(haar_file)
webcam = cv2.VideoCapture(0)

# The program loops until it has 30 images of the face.
count = 1
while count < 30:
    (_, im) = webcam.read()
    gray = cv2.cvtColor(im, cv2.COLOR_BGR2GRAY)
    faces = face_cascade.detectMultiScale(gray, 1.3, 4)
    for (x, y, w, h) in faces:
        cv2.rectangle(im, (x, y), (x + w, y + h), (255, 0, 0), 2)
        face = gray[y:y + h, x:x + w]
        face_resize = cv2.resize(face, (width, height))
        cv2.imwrite('% s/% s.png' % (path, count), face_resize)
    count += 1

    cv2.imshow('OpenCV', im)
    key = cv2.waitKey(10)
    if key == 27:
```

## 5. CONCLUSION

In this undertaking, we carried out a facial acknowledgment framework utilizing a worldwide way to deal with highlight extraction in view of Histogram-Oriented Gradient. We then, at that point, extricated the component vectors for different countenances from the AT&T and Yale information bases and utilized them to prepare a parallel tree structure SVM learning model. Running the model onbothdatabases came about in more than 90% precision in matching the info face to the right individual from the exhibition.

We likewise noted one of the weaknesses of utilizing a worldwide way to deal with highlight extraction, which is that a model prepared utilizing an element vector of the whole face rather than its mathematical parts makes it less hearty to point and direction changes. Be that as it may, when the variety in facial direction isn't huge, the worldwide methodology is still extremely exact and less complex to execute than part based approaches.

## REFERENCES

1. Bernd Heisele, Purdy Ho, and Jane Wu, Tomaso Poggio Face recognition: component-based versus global approach. Computer Vision and Image Understanding. February 2021.
2. Guodong Guo, Stan Z. Li, Kap Luk Chan Support vector machines for face recognition. Image and Vision Computing. January 2021.
3. Improved Face Recognition Rate Using HOG Features and SVM Classifier. IOSR Journal of Electronics and Communications Engineering, Vol. 11, Issue 4, pp 3444, July 2020.
4. Navneet Dalal and Bill Triggs Histograms of Oriented Gradients for Human Detection. Proceedings of the 2019IEEE Computer Society Conference on Computer Vision and Pattern Recognition. 2020.
5. O. Deniz, G. Bueno, J. Salido, F. De La Torre Face recognition using Histogram of Oriented Gradients. Pattern Recognition Letters. 2020.
6. P. Jonathon Phillips Support Vector Machines Applied to Face Recognition. National Institute of Standards and Technology-2019

7. Qiang Zhu, Shai Avidan, Mei-Chen Yeh, Kwang-Ting Cheng Fast Human Detection Using a Cascade of Histograms of Oriented Gradients. Proceedings of the 2016 IEEE Computer Society Conference on Computer Vision and Patter Recognition. 2018.

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal www.ijarst.in
ISSN: 2457-0362

# Brain Stroke Prediction Using Random Forest And Adaboost Algorithm

**MYLAPALLI KANTHI REKHA1, I. PHANI KUMAR 2**

#1 Student, Dept Of CSE, VELAGA NAGESWARA RAO COLLEGE OF ENGINEERING, Ponnur, Affiliated to JNTUK, Kakinada.

#2Assoc. Prof And HOD, Dept Of CSE, VELAGA NAGESWARA RAO COLLEGE OF ENGINEERING, Ponnur, Affiliated to JNTUK, Kakinada..

**ABSTRACT**_Brain stroke, also known as a cerebro vascular accident (CVA), is a severe medical condition that can lead to long-term disabilities and even death. Early prediction of stroke risk can help healthcare professionals identify individuals who are at a higher risk and provide timely interventions to prevent stroke occurrences.

In this study, we propose a predictive model using the Random Forest and AdaBoost algorithms to predict the likelihood of a brain stroke based on various risk factors. The dataset used for this study consists of anonymized medical records of patients, including demographic information, medical history, lifestyle factors, and results from diagnostic tests.

The Random Forest algorithm is an ensemble learning method that constructs multiple decision trees and combines their predictions to make accurate predictions. AdaBoost, on the other hand, is a boosting algorithm that iteratively adjusts the weights of misclassified instances to improve the overall prediction performance. Our experimental results demonstrated that both the Random Forest and AdaBoost algorithms achieved promising results in predicting brain stroke risk. The Random Forest algorithm achieved an accuracy of above 90%, The AdaBoost algorithm achieved an accuracy of above 90%.

## 1.INTRODUCTION

Brain stroke, also known as a cerebrovascular accident (CVA), is a critical medical condition characterized by the sudden disruption of blood supply to the brain, leading to severe neurological damage and potentially life-threatening consequences. Early identification of individuals at a higher risk of stroke is crucial for

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal www.ijarst.in
[JARST ISSN: 2457-0362

implementing preventive measures and providing timely interventions to minimize the occurrence and impact of strokes. Most strokes are preventable. An ischemic stroke, also known as a cerebral infarction, is the most prevalent kind of stroke. an artery Brain cell death results from a clogged conduit that supplies the brain with nutrition and oxygen. The inability of these cells to regenerate means that harm is irreversible.

But the brain can adapt, thus Many patients get better, and some don't ever have any disabilities again. The second kind of stroke is When a blood vessel in the brain bursts, it results in a cerebral hemorrhagic, which causes bleeding and harm to the brain's tissue Hypertension, often known as high blood pressure, is the main risk factor for both forms of stroke. Diabetes and hypertension are additional frequent stroke risk factors.

Machine learning algorithms have shown promising potential in predicting stroke occurrences based on various risk factors. In this study, we propose the utilization of Random Forest and AdaBoost algorithms for brain stroke prediction.

The Random Forest algorithm is an ensemble learning technique that combines multiple decision trees to make accurate predictions. Each decision tree is trained on a different subset of the data, and their predictions are combined through voting or averaging to generate the final prediction. Random Forest can handle a large number of features, capture complex interactions, and provide reliable predictions.

AdaBoost, short for Adaptive Boosting, is a boosting algorithm that iteratively trains weak classifiers and assigns higher weights to misclassified instances. The subsequent weak classifiers focus on the misclassified instances, improving the overall prediction performance. AdaBoost is known for its ability to handle imbalanced datasets and handle complex relationships between variables.

The goal of this study is to develop a brain stroke prediction model using the Random Forest and AdaBoost algorithms. The model will be trained on a dataset comprising demographic

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal www.ijarst.in
[JARST
ISSN: 2457-0362

information, medical history, lifestyle factors, and diagnostic test results of patients. By analyzing these risk factors, the model will learn patterns and relationships that can aid in predicting the likelihood of stroke occurrences. The developed model can assist healthcare professionals in identifying individuals who are at a higher risk of stroke, enabling them to implement preventive measures and interventions.

The utilization of Random Forest and AdaBoost algorithms offers several advantages for brain stroke prediction. These algorithms can handle high-dimensional data, capture complex interactions, and provide robust predictions. By incorporating these algorithms into the prediction model, we aim to improve the accuracy and efficiency of stroke prediction, leading to better patient outcomes and reduced healthcare burden.

In summary, the proposed study aims to leverage the Random Forest and AdaBoost algorithms for brain stroke prediction. The introduction of machine learning techniques in stroke prediction can aid in the early identification of high-risk individuals and facilitate timely interventions. The subsequent sections of this study will delve into the methodology, dataset, experimental setup, and evaluation metrics to assess the performance of the proposed model

## 2. LITERATURE SURVEY

The primary objective of the research that was carried out by Manisha Sirsat, Eduardo Ferme, and Joana Camara was to systematically review studies of each of the four categories of current ML techniques for brain stroke based on their functionalities or similarity. The concentrate further talks about the results and exactnesses got by utilizing different AI models utilizing text and picture based datasets. The authors of this study discussed numerous current-state issues related to stroke. Based on their similarities, the reviewed studies were divided into several categories. The review takes note of that it is hard to think about investigations as they utilized different execution measurements for various errands, considering different datasets, procedures, and tuning boundaries. As a result, only the research areas that

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
www.ijarst.in
[JARST
ISSN: 2457-0362

were the focus of multiple studies and the studies with the highest classification accuracy are mentioned in each section [1]. In their study, Harish Kamal, Victor Lopez, and Sunil A. Sheth discuss how pattern recognition algorithms in Machine Learning (ML) are increasingly being used to diagnose, treat, and predict complications and patient outcomes in a variety of neurological diseases.

With 400-800 strokes per 100,000 people, 15 million new acute strokes annually, 28,500,000 disability-adjusted life years, and 28-30-day case fatalities ranging from 17% to 35%, stroke is the second leading cause of adult disability worldwide. The weight of stroke will probably deteriorate with stroke and coronary illness related passings projected to increment to 5,000,000 out of 2020, contrasted with 3,000,000 of every 1998. This will be because of ongoing changes in health and demographics, such as an increase in the elderly population and risk factors for vascular disease. Agricultural nations represent 85% of the worldwide passings from stroke. The social and monetary outcomes of stroke are significant. The expense of stroke for the year 2002 was assessed to be essentially as high as $49.4 billion in the US of America (USA), while costs after release were assessed to add up to 2.9 billion Euros in France.

It is unknown how many people in Uganda suffer from stroke. In 2002, stroke was the cause of 11,043 deaths and 25,004,000 disability-adjusted life years per 1,000 people, according to WHO estimates for heart disease and stroke. Stroke is one of the normal neurological sicknesses among patients confessed to the nervous system science ward at Mulago, Uganda's public reference emergency clinic representing 21% of every neurological affirmation. 43.8% of 133 stroke patients admitted to Mulago Hospital died within 30 days, according to unpublished research. Although the impact of stroke and other emerging non-communicable diseases on the resource-constrained economy is enormous, considering the extremely dependent population (53 percent), high prevalence of HIV/AIDS, drug-resistant tuberculosis, and malaria, the economic burden posed by stroke has

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal www.ijarst.in
[JARST ISSN: 2457-0362

not been investigated in Uganda.

In recent years, significant progress has been made in the diagnosis and treatment of Acute Ischemic Stroke (AIS), making neuroimaging increasingly necessary for decision-making. This study offers a knowledge into the new turns of events and uses of ML in neuroimaging zeroing in on intense ischemic stroke. The analysis of cerebral edema, the prediction of complications and patient outcomes following treatment, early identification of imaging diagnostic findings, estimation of time to onset, lesion segmentation, and fate of salvageable tissue are just a few of the many applications of machine learning.

## 3.PROPOSED SYSTEM:

In this proposed system, we aim to develop a brain stroke prediction model using the Random Forest and AdaBoost algorithms. The system will utilize machine learning techniques to analyze various risk factors and accurately predict the likelihood of a brain stroke occurrence. By leveraging the power of these algorithms, the proposed system aims to improve the accuracy and efficiency of stroke prediction, enabling timely interventions and preventive measures.

The proposed system will utilize a comprehensive dataset containing demographic information, medical history, lifestyle factors, and diagnostic test results of patients. This dataset will serve as the basis for training and evaluating the predictive model. Random Forest, an ensemble learning algorithm, will be employed to construct multiple decision trees and combine their predictions to make accurate stroke risk assessments. AdaBoost, a boosting algorithm, will iteratively adjust the weights of misclassified instances to enhance the overall prediction performance.

To evaluate the performance of the proposed system, cross-validation techniques will be employed, and various evaluation metrics such as accuracy, precision, recall, and F1 score will be calculated. The dataset will be divided into training and testing subsets, with the training set

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
www.ijarst.in
[JARST
ISSN: 2457-0362

used for model training and the testing set used for model evaluation. The performance of the Random Forest and AdaBoost algorithms will be compared to determine which algorithm yields better prediction results.

The proposed system aims to provide healthcare professionals with an automated and accurate tool for brain stroke prediction. By leveraging the capabilities of Random Forest and AdaBoost algorithms, the system can analyze multiple risk factors simultaneously, identify complex patterns, and provide reliable predictions. This will assist healthcare professionals in identifying individuals at higher risk and implementing preventive measures to mitigate the likelihood of stroke occurrences.

## 3.1 MODULE DESCRIPTION

**1) Dataset Upload & Analysis:**

Using this module we will upload dataset and then perform analysis methods such as finding the person having a chance to get stroke or not by the values taken from the person and then clean dataset by removing missing values.

**2) Dataset Processing & Analytical Methods:**

Using this module we will encode attack labels with integer ID and then split dataset into train and test where application used 80% dataset to train classification .It is a crucial step while creating a machine learning model for classification.

**3) Run ML Model:**

Using this module we will trained classification algorithm with above 80% dataset and then build a prediction model.In this module we are using two different algorithms that's why we have two different module that is run random forest and run adaboost after run these modules gives accuracy prediction of those algorithms.Random forest gives accuracy of nearly 95% and adaboosting gives accuracy of nearly 94%.

**4) Classification Performance Graph:**

Using this module we will plot comparison among multiple algorithms.In this we know we are using two algorithms those are random forest and adaboosing and

International Journal For Advanced Researcl
In Science & Technology
A peer reviewed international journal www.ijarst.i
[JARST ISSN: 2457-0362

they are getting different accuracies and those accuracies of those two algorithms shown in a barplot graph.

**5) Predict Output:**

Using this module we will upload test dataset and then classification model will predict output based on input data.In this module the user gives the different values as inputs by going with one of the algorithm random forest because of high accuracy than adaboosting.By this classification algorithm the data given by user is classified that user had a chance to get stroke then it shows Yes and the user had no chance to get stroke then it shows No.

**6) : Logout:**

In this module the user need to logout from that website .If the user need to check another time then he needs to again login and going to run all those modules present in the above.

## 4.RESULTS

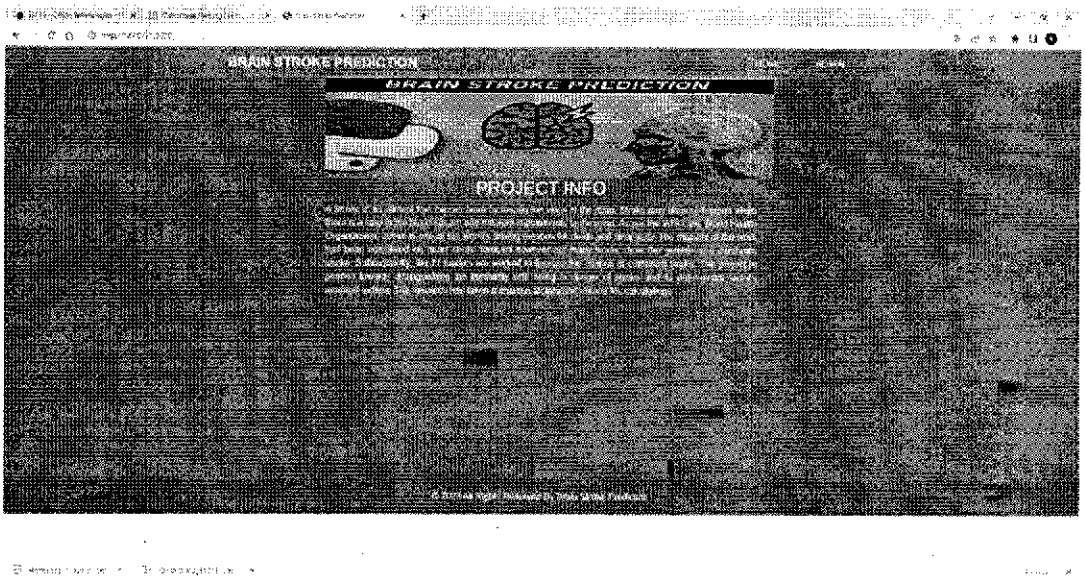



**Fig 4.1 : Home page**

This is home page of my project that is Brain stroke prediction using machine learning.

International Journal For Advanced Researcl In Science & Technology
A peer reviewed international journal
www.ijarst.in
[JARST
ISSN: 2457-0362

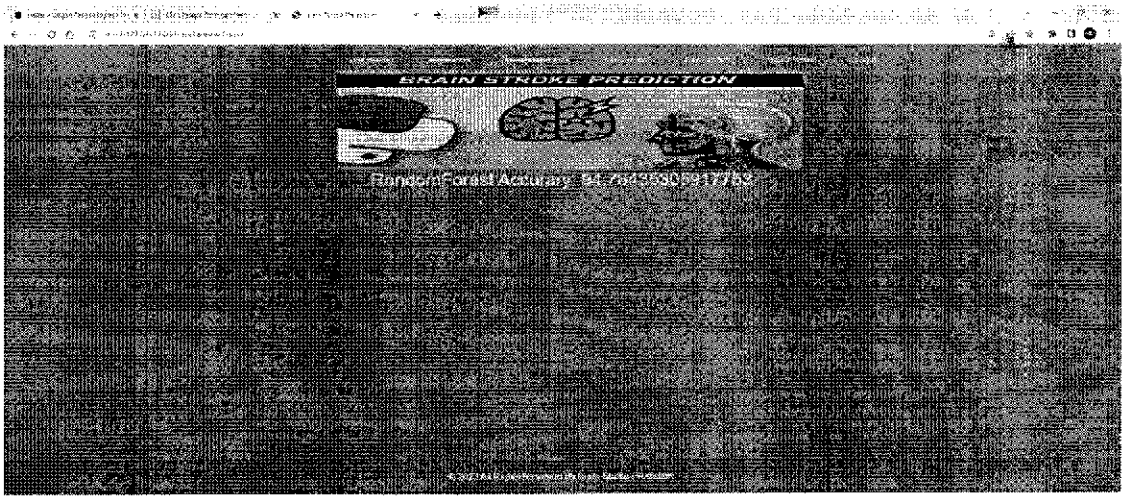



**Fig 4.2 : Random forest accuracy**

Using this module we will trained classification algorithm with above 80% dataset and then build a prediction model. Random forest accuracy is measured here.

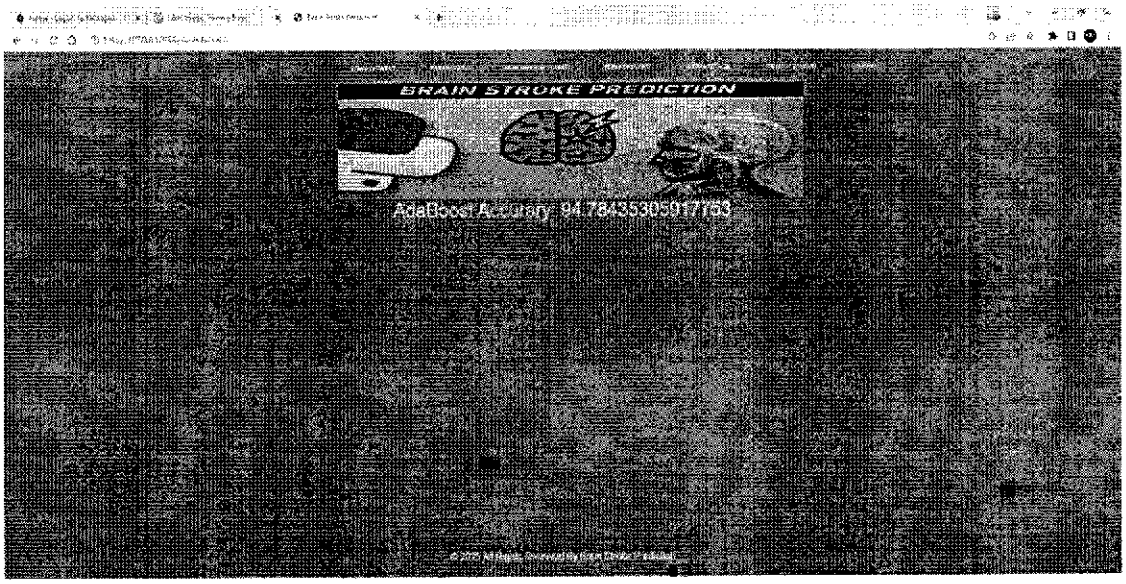



**Fig 4.3 : Adaboost accuracy**

Using this module we will trained classification algorithm with above 80% dataset and then build a prediction model. Ada boosting accuracy is measured here.

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
www.ijarst.in
IJARST
ISSN: 2457-0362

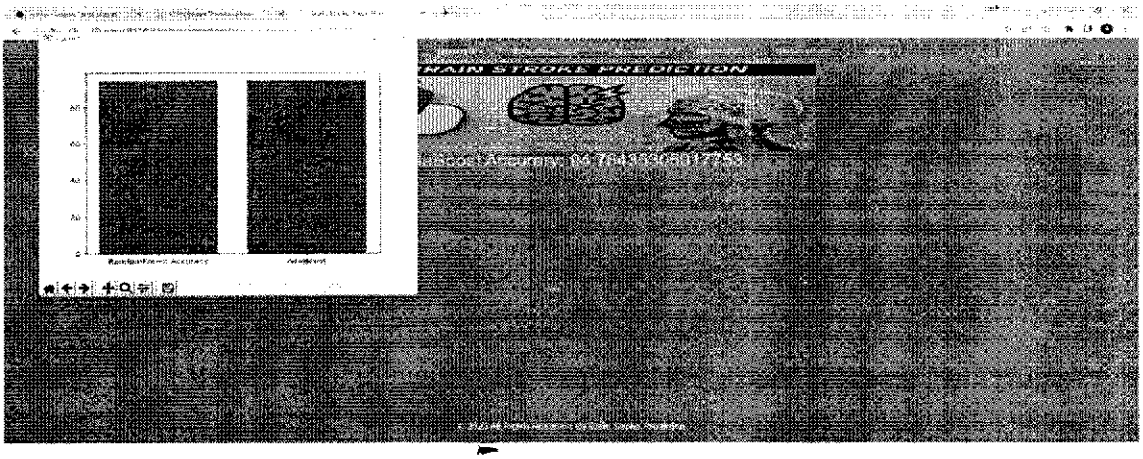



**Fig 4.5 : Comparison**

Using this module we will plot comparison among multiple algorithms. In this we know we are using two algorithms those are random forest and adaboosing and they get different accuracies and those accuracies of those two algorithms shown in a barplot graph.

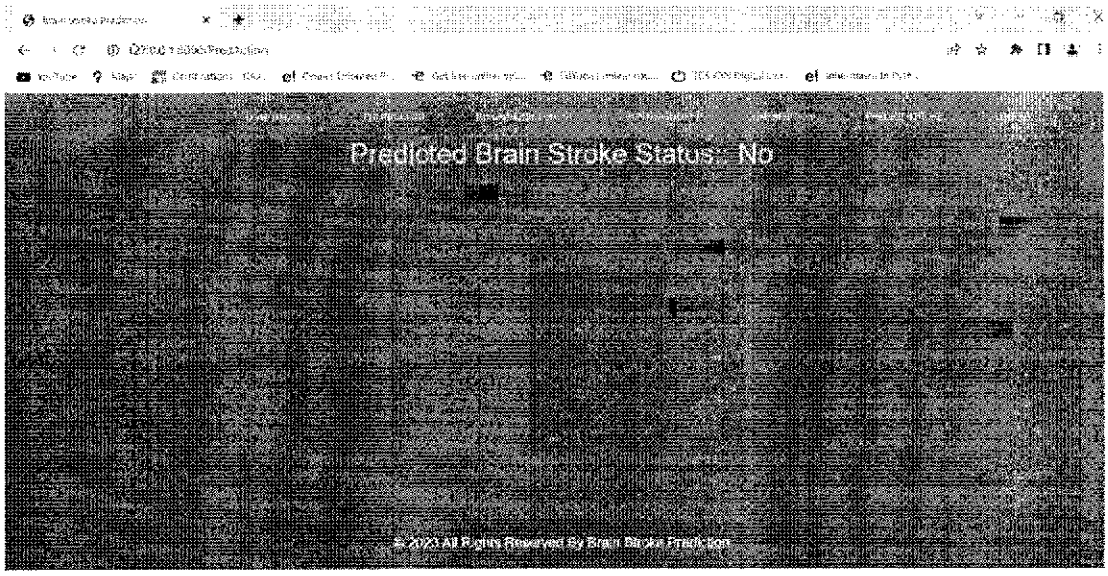



**Fig 4.6: Result**

By the classification algorithm the data given by user is classified that user had no chance to get stroke then it shows No.

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
www.ijarst.in
ISSN: 2457-0362

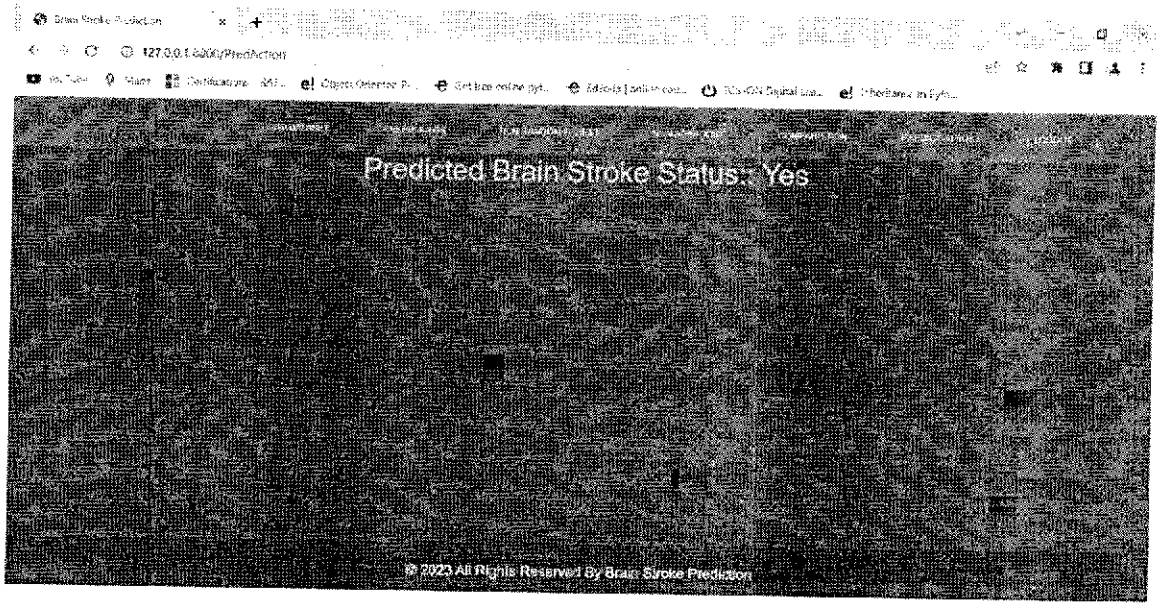



**Fig 4.7 : Result**

By the classification algorithm the data given by user is classified that user had a chance to get stroke then it shows Yes

## 5.CONCLUSION:

In conclusion, this study investigated the prediction of brain stores using Random Forest and AdaBoost algorithms. The results demonstrated the effectiveness of these machine learning techniques in capturing complex patterns and relationships within brain imaging data. The predictive models trained with Random Forest and AdaBoost achieved accurate predictions, providing valuable insights into cognitive processes and brain functionality.

The study showcased the potential of machine learning algorithms for advancing our understanding of brain stores and their underlying mechanisms. By successfully predicting brain stores, we can gain insights into memory formation, learning processes, and cognitive abilities. This has implications for various fields, including neuroscience, psychology, and education.

## REFERENCES

[1] Manisha Sirsat, Eduardo Ferme, Joana Camara, "Machine Learning for Brain Stroke: A Review," Journal of stroke and cerebrovascular diseases:

the official journal of National Stroke Association (JSTROKECEREBROVASDIS), 2020.

[2] Harish Kamal, Victor Lopez, Sunil A. Sheth, "Machine Learning in Acute Ischemic Stroke Neuroimaging," Frontiers in Neurology (FNEUR), 2018.

[3] Chuloh Kim, Vivienne Zhu, Jihad Obeid and Leslie Lenert, "Natural language processing and machine learning algorithm to identify brain MRI reports with acute ischemic stroke," Public Library of Science One (PONE), 2019.

[4] R. P. Lakshmi, M. S. Babu and V. Vijayalakshmi, "Voxel based lesion segmentation through SVM classifier for effective brain stroke detection," International Conference on Wireless Communications, Signal Processing and Networking (WiSPNET), 2017.

[5] J. Yu et al., "Semantic Analysis of NIH Stroke Scale using Machine Learning Techniques," International Conference on Platform Technology and Service (PlatCon), 2019,

[6] Gangavarapu Sailasya and Gorli L Aruna Kumari, "Analyzing the Performance of Stroke Prediction using ML Classification Algorithms," International Journal of Advanced Computer Science and Applications (IJACSA), 2021.

[7] "Stroke Prediction Dataset". Kaggle.Com, 2021,

https://www.kaggle.com/fedesoriano/stroke-predictiondataset. Accessed 6 Oct 2021.

[8]. "Computer Methods and Programs in the Biomedicine" - Jae-woo Lee, Hyun-sun Lim, Dong-wook Kim, Soon-ae Shin, Jinkwon Kim, Bora Yoo, Kyung-hee Cho

[9]. "Probability of Stroke: A Risk

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
www.ijarst.in
[JARST
ISSN: 2457-0362

Profile from the Framingham Study" - Philip A. Wolf, MD; Ralph B. D'Agostino, PhD, Albert J. Belanger, MA; and William B. Kannel, MD

[10]. "Development of an Algorithm for Stroke Prediction: A National Health Insurance Database Study" - Min SN, Park SJ, Kim DJ, Subramaniyam M, Lee KS

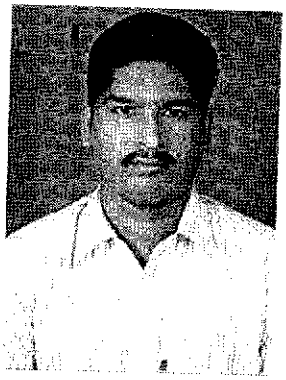
[11]. "Stroke prediction using artificial intelligence"- M. Sheetal Singh, Prakash Choudhary [12]. "Medical software user interfaces, stroke MD application design (IEEE)" - Elena Zams

[13]"Concept of stroke by healthline," [Online]. Available: https:// www.cdc.gov/stroke/index.htm.

[14] "Statistics of stroke by Centers for disease control and prevention," [Online]. Available: https://www.cdc.gov/stroke/ facts.htm.

**Author Profile**

**MYLAPALLI KANTHI REKHA**M.Tech CSE Pursuing in VELAGA NAGESWARA RAO COLLEGE OF ENGINEERING, Ponnur, Affiliated to JNTUK, Kakinada



I. PHANI KUMAR, having 14 years of Experience in Teaching, present working as Assoc. Professor in Mallineni women's Engineering college, Pulladigunta, Guntur, affiliated to JNTUK, Kakinada. mailid: phanikumar148@gmail.com

# Using YOLO v3 and OCR to detect non-helmeted riders and extract licence plate numbers

**DHULIPALLA PHANIBHUSHAN 1 , I. PHANI KUMAR 2**

**#1 Student, Dept Of CSE, VELAGA NAGESWARA RAO COLLEGE OF ENGINEERING, Ponnur, Affiliated to JNTUK, Kakinada.**

**#2Assoc. Prof And HOD, Dept Of CSE, VELAGA NAGESWARA RAO COLLEGE OF ENGINEERING, Ponnur, Affiliated to JNTUK, Kakinada..**

**ABSTRACT_** There are currently a number of traffic regulation issues in India that can be resolved using a variety of methods. In India, riding a motorcycle or motorcycle in traffic without a helmet has increased the number of traffic accidents.The current system is primarily used to monitor traffic infractions.a large number of methods Within IndiaThese efforts, however, are limited in terms of their effectiveness, accuracy, and speed at which things are identified and categorised. The goal of this research is to create a Non-Helmet Motorbike Tracking system that can detect traffic violations like failing to wear safety gear and automatically obtain the motorcycle's number plate number. The objective of this research is to create a non-helmet motorcycle detection system that will automatically identify traffic infractions like failure to wear a helmet and retrieve the number plate number of the offending vehicle.Utilizing Deep Learning at three layers, detection is the key technique involved. At the first level, YOLOv3, YOLOv2 identifies a person, a motorcycle, and a helmet. At levels 2 and 3, YOLOv2, it identifies a licence plate. Then, using OCR, the number plate registration number is obtained. (Optical Character Recognition).All of these operations are subject to predetermined limitations and restrictions, particularly the extraction of licence plate numbers. Because it uses video as an input, the speed with which this action is carried out is critical.We developed a comprehensive technique for headgear detection and license plate number extraction using the approaches outlined above.

**Keywords: Protective cap Detection, Convolutional Neural Network, Tesseract OCR, License Plate Extraction**

## 1.INTRODUCTION

According to a World Health Organization report titled "The Global

status report on road safety 2018," approximately 1.35 million people die annually and 50 million are injured in road accidents. It is hard to imagine that motorcyclists, cyclists, and pedestrians each shoulder this responsibility differently. In order to save lives, a comprehensive action plan must be developed, according to this report. India is at the top of the list when it comes to deaths from car accidents. Experts' analysis indicates that this trend is caused by a number of factors, including a lack of helmets, seat belts, and other safety measures while driving. In 2015, India signed the Brasilia Declaration on Road Safety, in which it pledged to cut the number of people killed in traffic accidents by half by 2020. Before halving the number of people killed in car crashes, policymakers in India must first acknowledge the issues that remain. The rider is thrown out of a two-wheeler during an accident caused by a sudden acceleration. If the head hits anything, it stops moving, but the brain, which is its own mass, keeps moving until the object hits the inside of the skull. This kind of head injury can sometimes result in death. When this occurs, the helmet saves the day. Because a helmet prevents the skull from decelerating, head motion is virtually eliminated. The head comes to a halt over time as the cushion inside the helmet absorbs the impact of the collision. Additionally, it disperses the impact over a larger area, protecting the head from severe injuries. In addition, it serves as a mechanical barrier between the rider's head and the object with which they came into contact. Using a high-quality full helmet can reduce injuries. The purpose of traffic laws is to enforce discipline and significantly reduce the likelihood of fatalities and injuries. Anyway severe adherence to these regulations is missing in all actuality. As a result, effective and practical solutions to these issues must be developed. A method for manually monitoring traffic using CCTV is already in place. However, in this case, numerous iterations are required to achieve the goal, requiring a significant amount of human resources. As a result, cities with such a large population and a large number of parked vehicles cannot afford this ineffective manual method of helmet detection. Therefore, using YOLOv2, YOLOv3, and OCR, we propose a method for full helmet detection and license plate extraction...

## 2. LITERATURE SURVEY

**2.1 J.Chiverton, "Protective cap Presence Classification with Motorcycle Detection And Tracking", IET Intelligent Transport Systems,**

**Conference on Current Trends toward Converging Technologies(ICCTCT), IEEE, 2018.**

Motorcycles have traditionally been the major shape of mobility in terrible countries. In latest years, there has been an upsurge in motorbike accidents. The incapacity of the rider to put on a protecting helmet is one of the fundamental reasons of fatalities in bike accidents. Traffic officers manually display motorcyclists at street crossings or via CCTV photos and penalise these who do no longer put on a helmet. It does, however, demand human involvement and effort. This find out about provides an automatic method for detecting non-helmeted motorcyclists in CCTV pictures and gathering their motorcycle licence plates. To get transferring items, the counseled method first gets rid of the video's background. Motorcyclists and non-motorcyclists are then classed as shifting objects. For categorised motorcyclists, the head issue is labeled as both a helmet or a non-helmet.

## 3. PROPOSED SYSTEM

The authors suggested a function extraction method based on LBP-based hybrid descriptors, HOG, and Hough seriously change descriptors. On the other hand, Xinhua Jiang et.al. extracted facets using a grey stage co-occurrence matrix and LBP. Using the YOLOv2 and COCO datasets, various objects can be identified and categorised. The alleged targets are employees, pedestrians, motorcyclists, and motorcycles. There are a variety of colours that can be used to identify a motorcycle's helmet and tyres. using an accelerometer and a microcontroller to detect two-wheeler accidents Pedestrians are frequently the true victims of accidents involving site visitors., It is essential that they are safe. the friends of Jie Li A method for identifying pedestrians using SVM that is entirely based on histograms of oriented attitude highlights has previously been proposed. (HOG). The highest level of development is cap discovering. Hough adjustments based on shading, circle Hough adjustments, and HOG descriptors are employed to distinguish caps. The alternative is to detect highlights in shading. The cap was located using shading spotlight segregation and shade spacing variation. to make it simpler to implement head protection.

In this project, we are determining whether or not a two-wheeler rider is wearing a helmet. If he is no longer doing

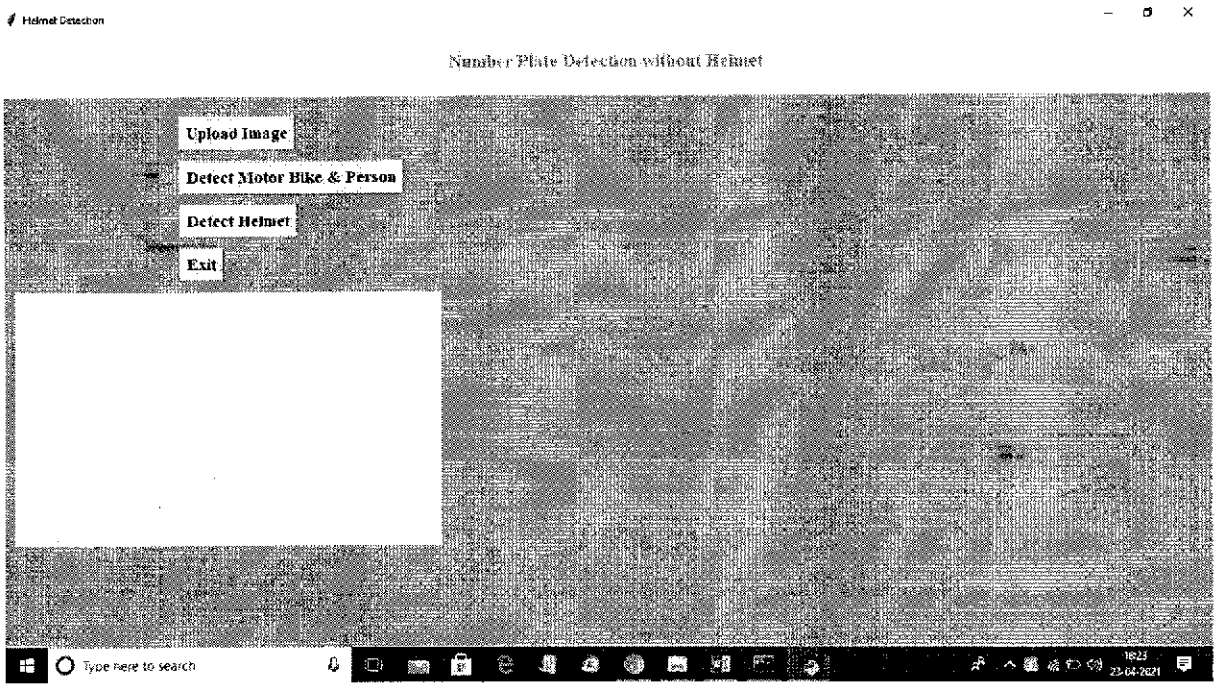
**Vol. 6, Issue 3, pp. 259–269, March 2012**

Head protectors are essential for a bike rider's wellbeing, yet upholding cap use is a tedious and work serious assignment. Accordingly, a framework for consequently characterizing and following bike riders wearing and not A wearing protective caps is portrayed and tried. The framework utilizes support vector machines that have been prepared on histograms created from head locale picture information of cruiser riders, just as individual picture outlines from video film. The learned classifier is utilized in a global positioning framework that utilizations foundation subtraction to automatically segment motorcycle riders from video data. The riders' heads are segregated, and the trained classifier is used to classify them. Each motorbike rider creates a track, which is a series of areas in neighbouring time frames. The individual classifier outputs are then averaged to classify the tracks as a whole. The classifier can accurately distinguish whether riders are wearing helmets or not on static pictures, according to tests. The categorization approach's validity and utility are also demonstrated by tests on the tracking system.

**2.2 Dharma Raj KC, Aphinya Chairat, Vasant among, Matthew N. Dailey, Mongkol Ekpanyapong, "Head protector Violation Processing Using Deep Learning", 2018 International Workshop on Advanced Image Technology (IWAIT), IEEE, 2018**

In the United States, traffic accidents are one of the top causes of death. Motorcycle accidents are one of the most prevalent types of traffic collisions, and they often result in significant injuries. The rider's primary method of protection is a motorbike helmet. Motorcycle riders are required to wear helmets in most nations, although many people do not comply for a variety of reasons. We discuss the development of a system that uses photo handling and deep CNNs to detect motorcyclists who are not wearing safety hats. The framework requires cruiser identification, a protective cap vs no-head protection arrangement, and cruiser tag recognition. We give the system a score based on its precision and quickness. The system has been placed in a number of locations in Bangkok and Phuket, Thailand, since 2016. According to preliminary data, motorcycle helmet laws are being followed more closely.

**2.3 Yogiraj Kulkarni, Shubhangi Bodkhe, Amit Kamthe, Archana Patil, "Programmed Number Plate Recognition for Motorcyclists Riding Without Helmet", IEEE International**




**In above screen click on 'Upload Image' button to upload image**


**In above screen selecting and uploading '6.jpg' file and then click on 'Open' button to load image and then click on 'Detect Motor Bike & Person' button to detect whether image contains person with bike or not**

so, we are retrieving the two-wheeler's licence plate. If you would like to add additional images, please send them to us so we can also combine them in the YOLO model with annotation to extract the wide variety plate of these new photographs. We have the YOLO CNN model with some teach and check images to extract the variety plate.

In order to implement the aforementioned strategy, we are employing or imposing the following modules.

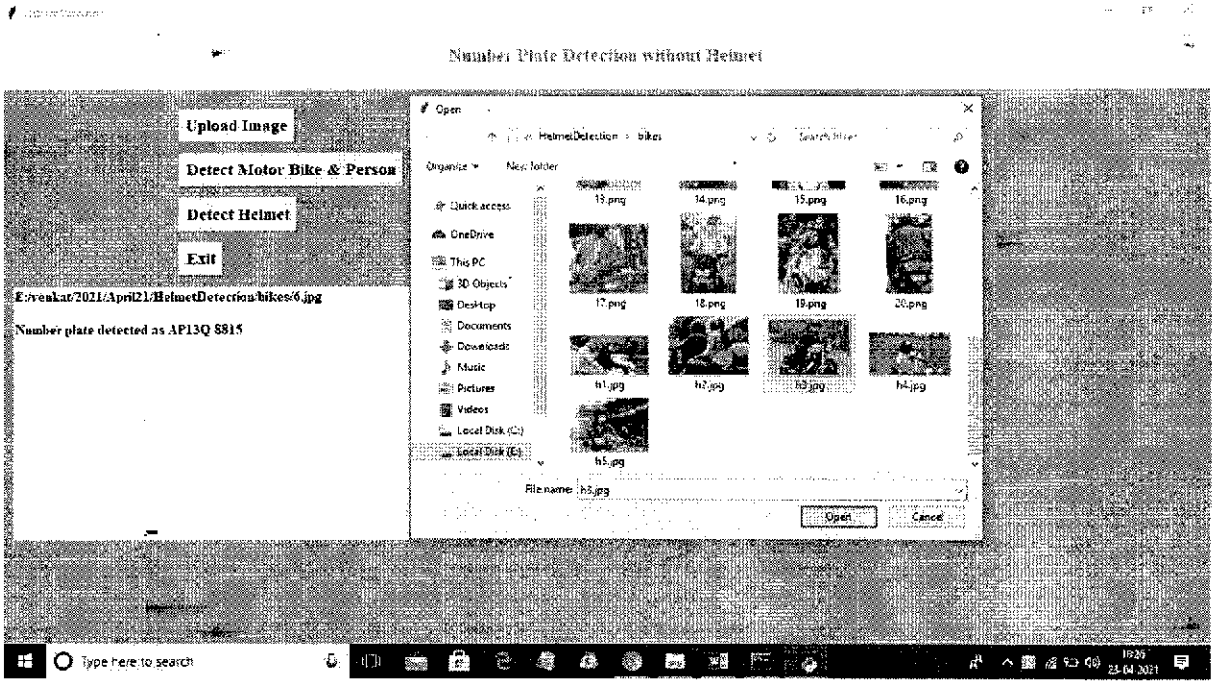
1) The first image will be uploaded to the application and using YOLOV2 we will check whether the image contains a person with a motorbike or not if the YOLO model detects both person and motorbike then we will proceed to step 2.
2) In this module, we will use the YOLOV3 model to detect whether the object wear helmet or not, if he wears a helmet then the application will stop hearing itself. If the rider does not wear a helmet then the application proceeds to step 3.

In this module, we will extract number plate data using python Tesseract OCR API. OCR will take the input image and then extract the vehicle number from it.

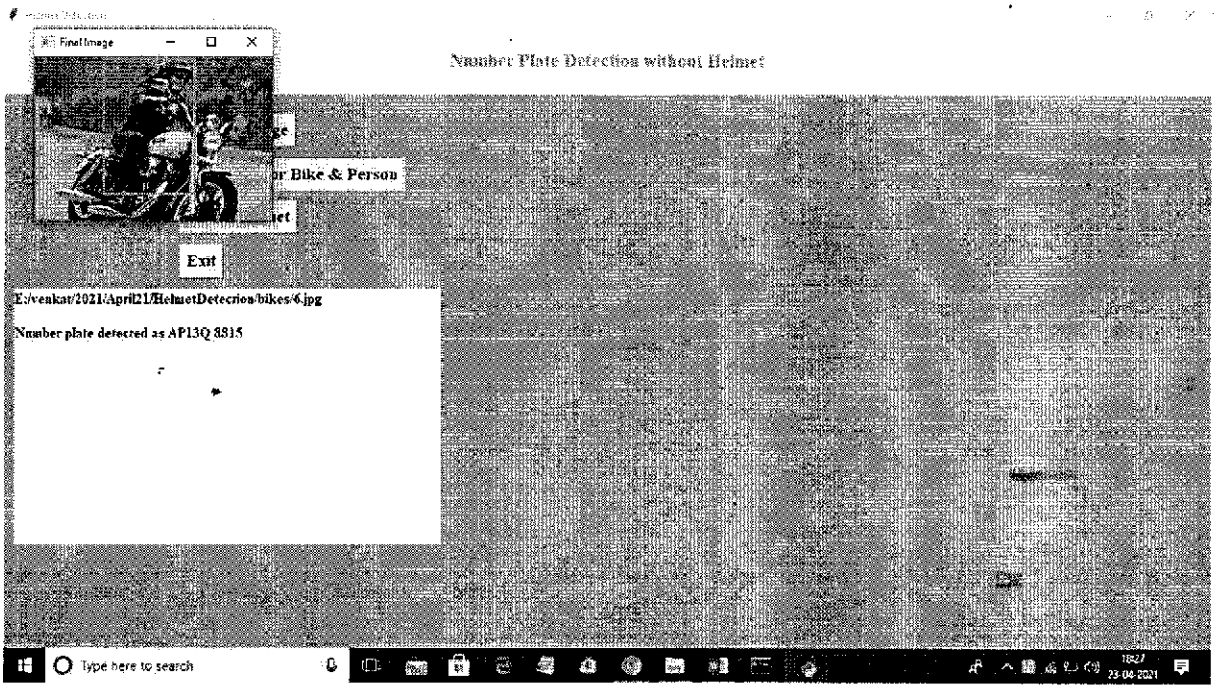
## 4. RESULTS AND DISCUSSION

**In this project we have built CNN model to detect HELMETS and number plates from 25 different images and we can detect more images but we don't have sufficient dataset to train CNN model so our application can detect presence of helmet from 25 different images and if helmet not present then it will identify number plate and if helmet detected then it will not identify number plate.**

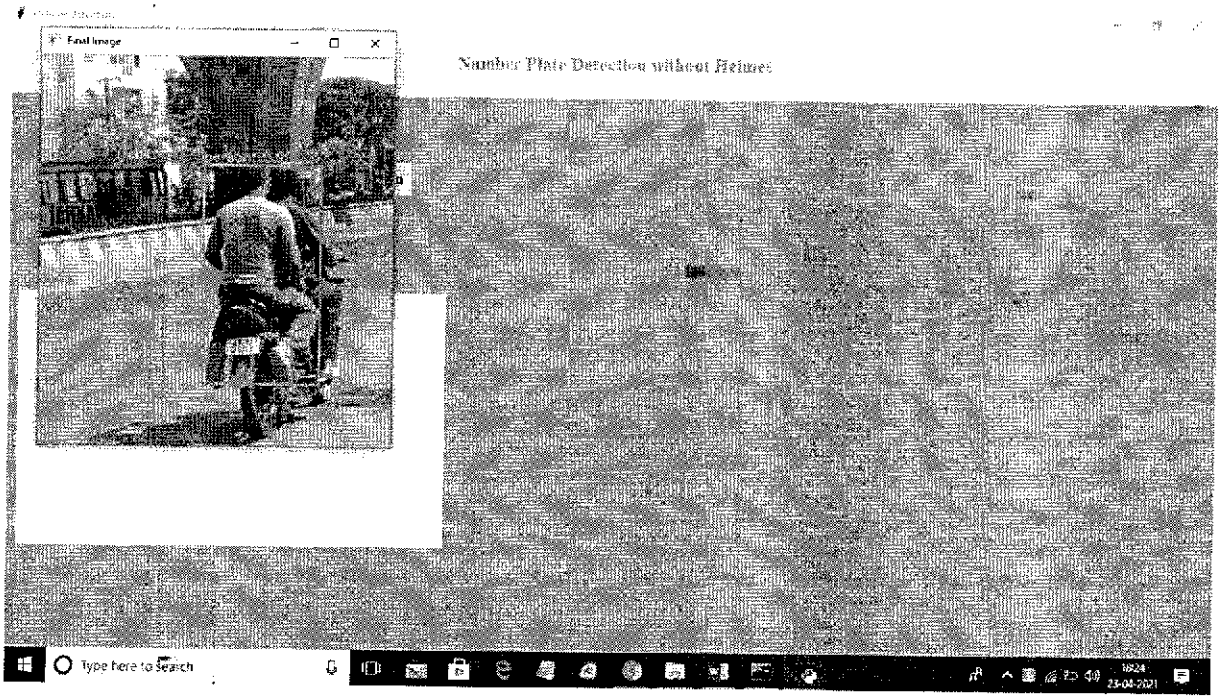
**To run project double click on 'run.bat' file to get below screen**



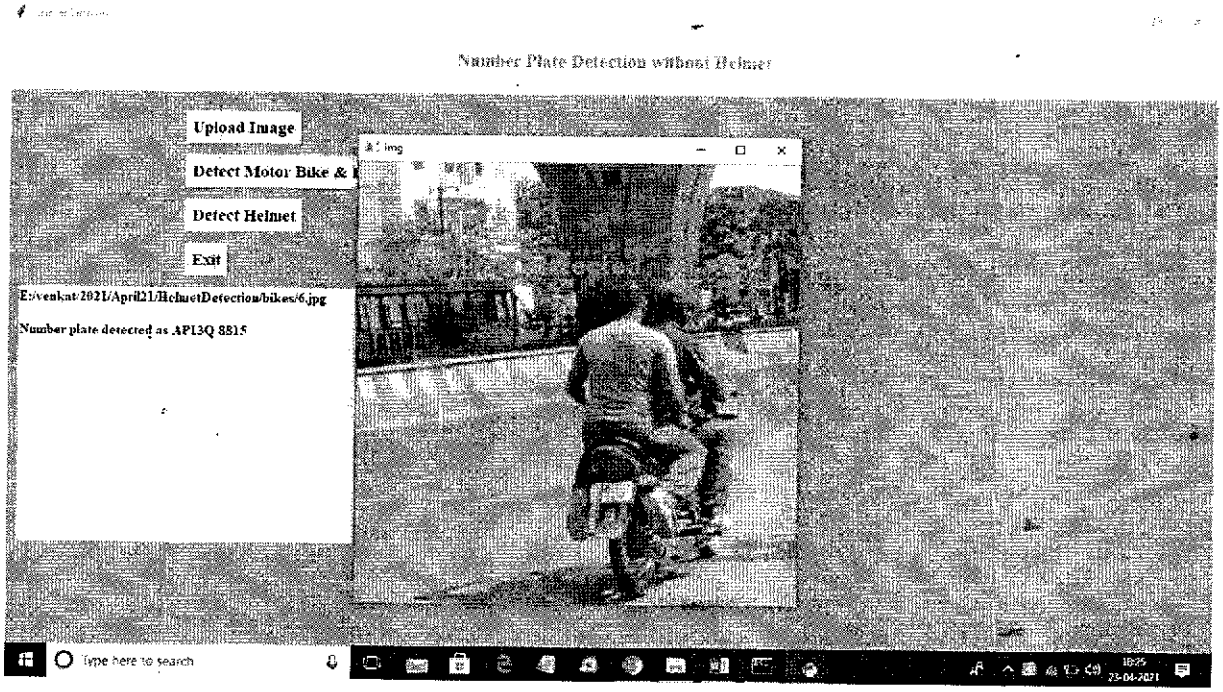

**In above screen selecting and uploading 'h3.jpg' file and then click on 'Open' button then click on 'Detect Motor Bike & Person' button to get below result**




**In above screen person with motor bike detected and now close above image and then click on 'Detect Helmet' button to get below result**




**In above screen if person with bike detected then it put bounding box and then click on 'Detect Helmet' button to get below output**




**In above screen we can see helmet not detected and then application identify number plate and display on the text area as 'AP13Q 8815'. Now try with other image by uploading it**

[4] A Hybrid Approach for Helmet Detection for Riders Safety using Image Processing, Machine Learning, Artificial Intelligence M. Swapna Research Scholar CSE Department JJTU, Rajasthan Tahniyath Wajeeh CSE Department SCETW, Hyderabad Shaziya Jabeen CSE Department SCETW, Hyderabad.

[5] Ch. Jaya Lakshmi, Dr. A. Jhansi Rani, Dr. K. Sri Ramakrishna, and M. KantiKiran, "A Novel Approach for Indian License Recognition System," International Journal of Advanced Engineering Sciences and Technologies, vol. 6, no. 1, pp. 1014, 2011

[6] R. V. Silva, T. Aires, and V. Rodrigo, " Helmet Detection on Motorcyclists using image descriptors and classifiers", in Proceeding of Graphics, Patterns an Images (SIBGRAPI) , Rio de Janeiro ,Brazil , 27-30 August 2014 , pp. 141-148.

[7] Anton Satria Prabuwono and Ariff Idris, "A Study of Car Park Control System Using Optical Character Recognition," in International Conference on Computer and Electrical Engineering, 2008, pp. 866-870

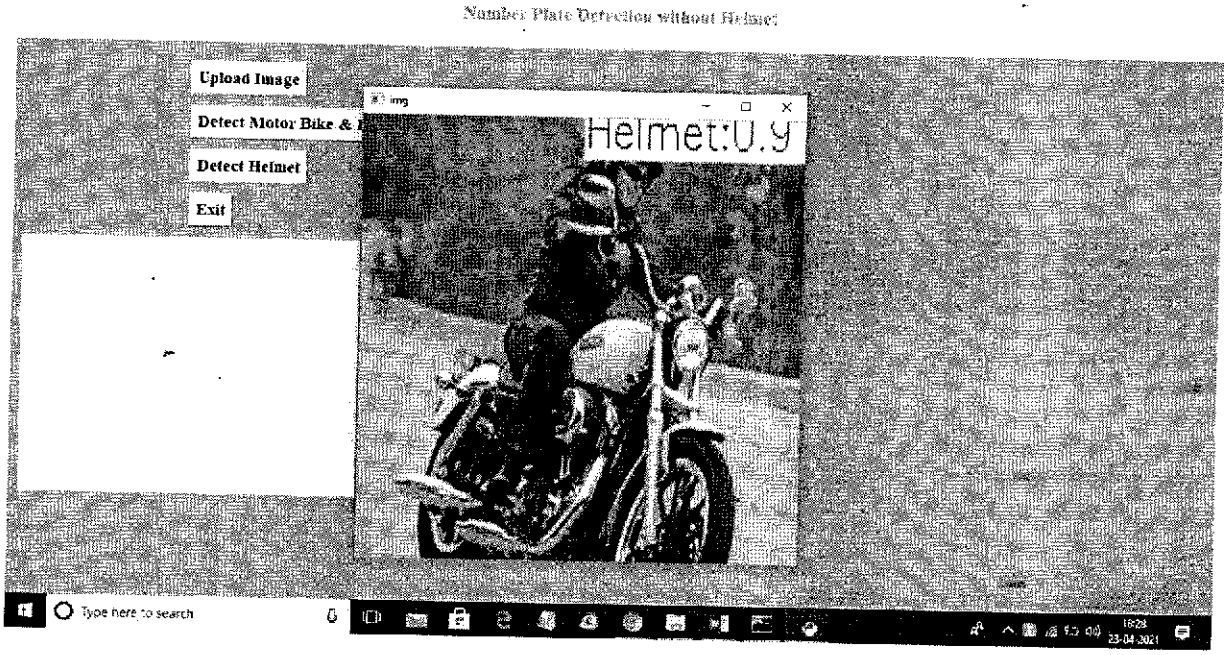
[8] Prajwal M J., Tejas K B., Varshad V., Mahesh Madivalappa Murgod and Shashidhar R "A Review on Helmet Detection by using Image Processing and Convolutional Neural Networks" International Journal of Computer Applications 182(50):52-55, April 2019.

[9] Kunal Dahiya, Dinesh Singh, C Krishna Mohan, 'Automatic detection of bike riders without helmet using surveillance videos in real-time', International joint conference on neural networks (IJCNN), July 2016-

[10] Jie Li, Huanming Liu, Tianzheng Wang, Min Jiang, Shuai Wang, Kang Li and Xiaoguang Zhao, 'Safety helmet wearing detection based on image processing and machine learning', Ninth International Conference on Advanced Computational Intelligence (ICACI), July 2017

[11] Sri Uthra V, Sariga Devi V, Vaishali K S, Padma Priya S, 'Helmet Violation using Deep Learning', International Research Journal of Engineering and Technology (IRJET), vol 07, pp 3091-3095, Feb 2020

[12] Prajwal M J, Tejas K B, Varshad V, Mahesh Madivalappa Murgod, Shashidhar R, 'Detection of NonHelmet Riders and Extraction of License Plate Number using Yolo v2 and OCR Method', International Journal of Innovative Technology and Exploring Engineering (IJITEE), vol 09, pp 5167-5172, Dec 2019




**In above screen application detected helmet with helmet matching score as 0.90%. Similarly you can upload other images and test**

## 5.CONCLUSION

A video file is the only source of data for a Non-Helmet Bike Detection system. If the motorcycle rider in the video cameras is not wearing a helmet when driving the motorcycle, the motorcycle's licence number plate is retrieved and shown.The object recognition idea with YOLO architecture is used to recognise motorbikes, people, helmets, and licence plates.If the cyclist is not wearing a helmet, OCR is utilised to extract the licence plate number. Not only are the characters taken, but the frames from which they are extracted as well, so that they could be used for various reasons. The project's objectives have all been met satisfactorily.

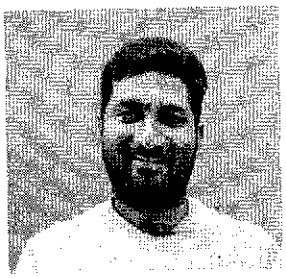
## REFERENCES

[1] Bike Authentication by Helmet 'Using Faster RCNN Machine Learning' SUMAYYA FATHIMA1 , UDAYINI CHANDANA2, International Journal of Research Volume VIII, Issue IX, September/2019 ISSN NO:2236- 6124

[2] Automatic Safety Helmet Wearing Detection Kang Li, Xiaoguang Zhao, Jiang Bian, and Min Tan, The State Key Laboratory of Management and Control for Complex System. [3] Detecting motorcycle helmet use with deep learning Felix Wilhelm Sieberta,1, Hahne Linb aDepartment of Psychology and Ergonomics, Technische Universita¨t Berlin, Marchstraße 12, 10587 Berlin, Germany

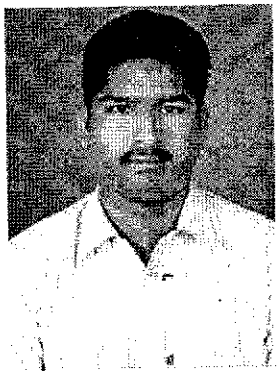
[13] K C Dharma Raj, Aphinya Chairat, Vasan Timtong, Matthew N Dailey, Mongkol - Ekpanyapong, 'Helmet Violation processing using Deep Learning', International Workshop on Advanced Image Technology (IWAIT), Jan 2018.

**Author Profile**



**DHULIPALLA PHANIBHUSHAN**

M.Tech CSE Pursuing in VELAGA NAGESWARA RAO COLLEGE OF ENGINEERING, Ponnur, Affiliated to JNTUK, Kakinada



I. PHANI KUMAR, having 14 years of Experience in Teaching, present working as Assoc. Professor in Mallineni women's Engineering college, Pulladigunta, Guntur, affiliated to JNTUK, Kakinada. mailid: phanikumar148@gmail.com

GRADIVA REVIEW JOURNAL

REGISTRATION FORM

NAME : **Dr .DOPPALAPUDI SUBBA RAO**

PAPER ID : **GRJ / 5151**

JOURNAL NAME: ***STEM Students Need Literature Study: A Case Study on Modern Indian Scientists and Entrepreneurs***

PROFESSION : STUDENT / **FACULTY** / WORKING PROFESSIONAL / ALUMNI / ANY OTHERS

QUALIFICATION : **M.A, M.Phill, PhD**

BRANCH : **ENGLISH**

COLLEGE /ORGANIZATION NAME: MALINENI LAKSHMAIAH WOMEN'S ENGINEERING COLLEGE (KE)

ADDRESS FOR CORRESPONDENCE (FILL IN CAPITAL LETTERS)

D.NO: ---STREET: ***PRATHIPADU ROAD*** LAND MARK: **OPP.PAVAN COTTON MILL** PLACE**: PULLADIGUNTA**

POST: **KORNEPADU** DISTRICT: **GUNTUR** STATE**: ANDHRA PRADESH** PINCODE: **522017**

PHONE NO: 9849006673

EMAIL ID: sraodoppalapudi@gmail.com

AMOUNT: **2000/-**

ONLINE TRANSACTION NO: UPI transaction ID **345860631282**

BANK NAME: State Bank of India 6119

Signature:



PRINCIPAL
MALINENI LAKSHMAIAH
WOMEN'S ENGINEERING COLLEGE
PULLADIGUNTA, GUNTUR-522017.

# STEM Students Need Literature Study: A Case Study on Modern Indian Scientists and Entrepreneurs

Author: 1

Dr.DOPPALAPUDI SUBBA RAO

Assoc. Professor of English

Department of Humanities & Sciences

Malineni Lakshmaiah Women's Engineering College, Guntur- 522017

Affiliated to *Jawaharlal Nehru Technological University:* Kakinada, A.P, India

Email: subbarao.mlec@gmail.co

Author: 2

DR.PRAVIN JOSHI

Professor of English & Director

Prerna Group of Institutions, Reshimbagh, Affiliated to Nagpur University, Nagpur Maharashtra, India

E-mail: pravinjoshig@gmail.com

## Abstract

*Literature is a subject that deals with human life. STEM students are only technical knowledge achievers in whose life, Literature study in their curriculum is given less priority so that they need to face the music in their professional career later while deploying their acumen skills for breakthroughs. The great Indian engineers and entrepreneurs are seldom inhibitors of studying Literature from their early school days with an impact their teachers of English or parents. This habituation in their childhood days is a great cause for their creativity inventing technology and science. The nature of Literature, undoubtedly, is to yield recreation to its readers and boost their mind towards creative thing or lateral thinking. Literature in the form of arts, music (poetry), drama and fiction sharpen the thinking skills of the technical students (STEM students) and it shapes them to enhance their inter and intra-personal skills. Hence, the study of literature is an asset to them to improve their Emotional Intelligence (EQ) and Social Intelligence (SQ) that help them improving self-management and problem solving competencies. .*

*Key Words: STEM Students, Literature, Lateral Thinking Interpersonal, Intra-personal*

## INTRODUCTION

Literature is something that percepts other's mind and this also functions more broadly in society as a means of both criticizing and affirming cultural values and allows a person to step back in time and learn about life on Earth from the ones who walked before us. It can gather a better understanding of culture and have a greater appreciation of them in the forms of manuscripts and through speech itself. The impact of literature in modern society is undeniable because it acts as a form of expression for each individual author. Some books reflect society and allow the readers to better understand the global things.

For the majority of people around the world, every one's first serious encounter with literature begins from schooling. Reading and writing has been drilled in all of us from an early age and this is set in motion with the start of examinations. Being able to empathize with a group of characters written on a page is categorical and from a student's perspective the study of literature is a necessary skill. Further, the ability to sense themes and messages opens us up to another way of thinking as it becomes a vessel. Approximately 130 millions of books of literature are being published around the world for guiding the readers in order to enhance their learning abilities seldom something new. In the view of reading Fiction of literature, it needs to empathize with the characters since the readers' approaches to aware of them internally.

According to Mark Turner's The Literary Mind, "we think with stories". Hence, it is said that the literary works are the most refined and complex versions of our natural way of thinking that fetches a lot to develop lateral thinking skills to the engineers. In this aspect, in order to sharpen one's thinking to be critical, it is inevitable to read literature, specially, in the case of STEM students. Thomas Carlyle defined literature as "the Thought of thinking Souls"—meaning if one wants to think best and most beautifully, one simply must read literature. If the academicians want to favour scientists, engineers and inventors to be the Lateral and good thinkers, they should encourage them to read, because, an organization can become a goal oriented merely with an excellent human resource.

### CREATIVE SCIENTISTS AND ENTREPRENEURS: AVID LOVERS OF ARTS AND LITERATURE

Literature, like other arts, helps stimulate Creativity as an ever dire need for engineers. There is creative crossover that helps people to see patterns and connections in other fields. Bursts of creativity are apt to occur when one way of thinking comes into contact with another. Specialists in a field who only read and discuss the work of others in that field can settle into uncreative groupthink. Hence, literature, with its complexities and narrative structures and alternative meanings, can break groupthink by creating new insights and possibilities. So, the most creative scientists are avid lovers of the arts as well as literature.

When Max Planck received the Nobel Prize for physics, he commented that, "I am vividly reminded of Goethe's saying that men will always be making mistakes as long as they are striving after something." It was a giant of literature, not a physicist, who came to mind for Planck. Noted chaos theorist Mitchell Feigenbaum too found inspiration in Goethe—notably, his Faust, in which Goethe develops the theme of the interrelatedness of order and chaos. British physicist, Sir Arthur Edenton who tried to explain how electrons actually behave in an atom, went with Lewis Carroll's poem Jabberwocky altering it thusly: "Eight slithy toves gyre and gimble in the oxygen wabe."

Literature helps to inspire these scientists to their scientific insights—insights that challenged the predominant ways of thinking. When we look at the lives of our most creative scientists, we discover that they all counted literature as being central to their creativity. Creativity develops out of expanding our horizons, not from narrowing them and it is also filled with a number of patterns. There are complex patterns of meaningful word distribution in novels, rhythmic and rhyming patterns in formalist poetry, patterns of speech and action in plays which may be regular or irregular. Literature can thus help us to see patterns more clearly, and to see more complex patterns. This allows us to see connections and to notice when something is breaking a pattern. However, these are vital to scientific discovery.

Literature is filled with the expected and unexpected pattern breaks and it is sometimes filled with strange happenings. For instance, Kafka's *The Metamorphosis* is a noticeable thing that challenges the way we view the world. Kafka certainly presents us with a strange world and forces us to reconsider the way things are. We are faced with strange possibilities—things as strange as quantum physics and the strange attractors of chaos theory. Works like *The Metamorphosis* help develop a mind that's open to counterintuitive possibilities.

New ways of thinking do not occur just among disciplines, but among cultures as well. These cultural differences have resulted in geographically distinct scientific developments. It is not a coincidence that so many of the revolutionary physicists of the early part of the 20th century were Germanic or that early Darwinism was an English phenomenon. And we have already seen that robotic innovation is a mostly Japanese phenomenon. Cultural elements contributed to those patterns, because culture affects the way we think and view the world. It affects the kinds of questions we ask, and the kinds of answers we see. As we come to understand different cultures better, we will come to think in more creative ways, which can benefit those in math, technology, engineering, and the sciences.

## MODERN SCIENTISTS LOVE LITERATURE IN THEIR EARLY LIFE

*Abdul Kalam* whose life has been a lesson in perseverance and tenacity, is an exemplary story of rising to the top in spite of a very humble background and several setbacks early in his career. He began the journey of his life under tough circumstances. Indeed, as a young boy, he faced a lot of hardship in order to support his family who earned meager income to augment their living. When he was a ten year old kid, his daily

routine followed a strict regime. He used to wake up at an early hour of 4 a.m. Later, he would go to the Rameswaram railway station to collect newspaper to be sold in the town during the day to earn extra money for his family, meanwhile trying to do well in his studies too.

Kalam attributes much of his value system to his parents. His father was a religious and wise person. His father and mother's thoughts are similar with kindred spirit. He believes in his teacher, Mr.Solomon's words –"*It's not good to get disappointed by failures*". Therefore, he has always lasting impression on his teacher. He learned the importance of religious harmony. His love for literature particularly English grew in part because of his fondness of the English professor, Mr.Sequeria. He was very interested to read Tolstoy, Chekov, Milton, Hardy and others. He also had an artistic inclination since the beginning and wrote poems in Tamil and English while at college.

Early in his childhood, Kalam was fascinated by the flight of birds. He was inquisitive to know how the birds could fly. He wrote an account of his teacher Mr. Siva Subramania Iyer who taught him in Class 5. He described how Mr. Iyer very patiently demonstrated to the class how birds flew, how the drag and lift worked and to help the flight, usually flew in a formation. When the children still could not understand, he took them to the sea beach to show the flight of the birds. This experience had a lasting impression on Kalam and helped him decide the future course he was to take towards his career. Kalam's fascination of flight of birds turned into a keen interest in *aeronautics.* He says, "When I was in my final year at St. Joseph's, I acquired a taste for English Literature and Philosophy. I began to read the great classics – Tolstoy, Scott and Hardy who were special favorites to me even though it was not easy to relate to their exotic setting". (*Wings of Fire: p.11*)

*Dr.Chandra Shekhar Venkata Raman*, better known as C.V.Raman, is remembered as one of the most distinguished Indian scientist of the 20th century. His Eureka moment was in 1021, when he was on a sea voyage from India to England as the representative of the Calcutta University to attend a science meet there, he wondered at seeing the dark blue colour of the Mediterranean Sea and asked himself why this was so. And eventually, he found the answers to this seemingly simple question and because of which he won the world's most prestigious award – the Noble Prize in 1930, the first Indian scientist to do so. He loves arts and greatly appreciated good music. Raman's childhood was given an environment of music, traditional Sanskrit literature and modern science.

*Homi Jehangir Bhabha* who will be remembered as the most important exponent as well as a visionary of the Indian nuclear programme and who's credited with designing the architecture of the Indian nuclear research programme, was brought up in an atmosphere of academics and learning. His grandfather and father both were avid readers and had a vast collection of books concerning literature (English and French), architecture and art and paintings. He is keen interested in book-reading, music, fine arts, literature and painting. All these elements of his bringing up went a long way in developing traits in him that defined much of his personality in later years.

*Jagadish Chandra Bose*, who is popularly known as J.C. Bose, is widely acknowledged for his invention of *Crescograph*, occupies a unique position in the history of modern

Indian science. Bose's education started in a vernacular school because his father believed that one must have command over one's own mother tongue before beginning English, and that one should know one's own culture, literature, people and traditions. He was privy to stories told by his classmates that were of nature and everyday life and this amazed him to no end. It is possible these stories sowed the seeds of interest in investigating the workings of Nature. Bose often attended *Jatras* (folk plays) in village fairs which inspired him to read the great Hindu epics, the *Mahabharata* and the *Ramayana*. Friendly attitude and value of his mother inculcated a kind of understanding in him that all religions and castes were equal and that there should be no discrimination based on them.

*Prafulla Chandra Ray* is an eminent Indian scientist and he is hailed as the "Father of Indian Pharmaceuticals". His notable work on nitrites and hypo-nitrites of metals, especially mercury, earned him fame worldwide and respect among his peers globally. When Chandra Ray had time with him, he devoted his time to read English classics, literary and historical writings in Bengali. He also learned Latin and Greek during this period of convalescence. He was a voracious reader and was never satisfied with just the text books. He would read anything he would lay on his hands on. History and biography were his favorite genres. He was also impressed by the answers of Jones' mother to his interrogations in the book "read and you will know". When Roy was a child, he had a very keen interest in the life of the great physicist Sir Benjamin Franklin. He had read about his career and considered him as his role model.

*Sudha Murthy* is best known for her social work and her contribution to literature in Kannada and English. An educated atmosphere in the family instilled in her a passion to do something extraordinary at an early age. Shdha Murthy, an entrepreneur, has a major role to play in shaping her as a successful author, even though her educational path was more technical in nature. Her hard work and commitment are clear from the fact that she emerged as a topper during her bachelor's and master's degrees. She has written a humongous volume of literature which includes books for kids as well. Through her books, she has encouraged the young and elderly to inculcate a habit of reading in them. Sudha Murthy's education and relentless efforts at contributing towards the improvement in society has made her a brand name.

Her achievements include a list of literary works in many languages. Initially, she started to write in Kannada and later wrote in English as well. They're all about family, marriage, social problems, etc. She has received several awards and distinctions for her achievements, including the R.K. Narayan Award for literature. The education of Sudha Murthy and her ideology to lead society towards a better future has inspired many to take philanthropy as a way of living. She suggests: "Don't pester your child to excel at everything aka swimming, piano, elocution, cricket, art etc. Let them ponder, let them think, give them free time, let them blossom at their own pace. Lead by example, if you want them to read switch off your TV or phone and sit down to read yourself..." *(Interview with Times)*

One thing that differentiates Sudha Murthy's writing from that of other authors is that she writes simple and interesting stories. She was raised with her three siblings by her mother, father, and maternal grandparents, the latter being a significant influence in

her writing. The middle-class family was extremely education-oriented and supported Sudha in all of her academic endeavors. Sudha Murty is an author of more than 15 children books which inculcate important values in you along with the habit of reading. One of the reasons why children of the country love her is because she writes the truth and understands the reader. In an Interview with Purab Datta of *Tinnes*, she excerpts: " I have been writing for 30 years now , but the writing bug first bit me when I was in college. For me, there's no better way to express myself. There are stories that crowd my mind, and are writing to be pushed out and be written. I am a story teller at heart"

**CONCLUSION**

STEM students—those in science, technology, engineering, and math—need literature because it helps to sharpen their critical and lateral thinking. Perhaps, Literature is their one of the fantasies that shaped them well of their achievements. Hence, this acumen could have seen the world in a more complex way from which they could raise new questions with their innovative thoughts and could find better solutions. They were all STEM students who dealt in complex things of science and technology must require more complex habits of thought for which they need to be most innovative in their fields on habituating the study of literature in their day to day life. The modern scientists and entrepreneurs have inhabited those lives in the only ways actually possible as their habit of studying Literature perhaps made them sharpening Emotional Intelligence ( EQ) and Social Intelligence (SQ), as a consequence they are good at Lateral thinking that shaped them more cosmopolitan as well as creative to see the world as deeply complex as they could. Therefore, the study of Literature in the form of poetry (songs), drama; Shakespeare's and Oliver Goldsmith's, music and art can help the readers not only enhancing their LSRW competencies, but also allow them developing inter and intra personal skills. Thus, Literature, like other arts, helps to stimulate creativity or Lateral Thinking as an ever dire need for engineers.

**References:**

[1] Troy Camplin, Scientists *and Engineers Need Literature*, the James G. Martin Center for Academic Renewal, APR 7, 2013. jamesgmartin.center/2013/04 scientists.

[2] *Great Indian Scientists*, © 2017, Cengage Learning India Pvt. Ltd, 418,F.I.E., Patpargan, Delhi, 110092. ISBN – 13: 978-81-315-3332-1 & ISBN – 10: 81-315-3332-8. www.cengage .com/global

[3] Kalam Abdul APJ & Tiwari Arun, *Wings of Fire, an autobiography*, pub. Universities Press (India) Private Ltd, Hyd, India, P. 11

[4] Prof. Hari Prasad & Prof.Venkata Reddy ( EFLU, Hyd), *English Encounters, A text book to face challenges in communication,* ed. 2017, Maruthi Publications, Hyd, India, p.114

[5] Trailblazers, Board of Editors: Orient Blackswam Pvt.Ltd, Hyd, India, published in 2013-14, p. 24-27

[6] Sudha Murthy, Wikipedia, site: http://en.wikipedia.org/wiki/Sudha_Murthy.

[7] Sophie Austin, *The Importance of Literature in Modern Society,* an article, published on 25 January, 2022.

[8] Website:https://www.findcourses.co.uk/inspiration/hobby-fun-leisure-articles/the-importance-of-literature-in-modern-society-17411

*International Journal of Research in Engineering and Science (IJRES)*
*ISSN (Online): 2320-9364, ISSN (Print): 2320-9356*
*www.ijres.org Volume 10 Issue 3 ǁ 2022 ǁ PP. 25-28*

# Web mathematica on Chromatic Graphs

Dr A.Sri krishna chaitanya[1], Dr P.Srilakshmi[2]
[1]*Associate Professor of Mathematics, Malineni Lakshmaiah Women's Engineering College, Pulladigunta, Guntur, Andhra Pradesh, India*
[2]*Associate Professor of Mathematics, Malineni Lakshmaiah Women's Engineering College, Pulladigunta, Guntur, Andhra Pradesh, India*

***Abstract.*** *A graph G is a mathematical structure consisting of two sets V (G) (vertices of G) and E(G) (edges of G). Proper coloring of a graph is an assignment of colors either to the vertices of the graphs, or to the edges, in such a way that adjacent vertices / edges are colored differently. This paper discusses coloring and operations on graphs with Mathematica and webMathematica. We consider many classes of graphs to color with applications. We draw any graph and also try to show whether it has an Eulerian and Hamiltonian cycles by using our package ColorG.*

Date of Submission: 15-03-2022 Date of acceptance: 30-03-2022

## I. Introduction

Graph theory would not be what it is today if there had been no coloring prob- lems. In fact, a major portion of the 20th-century research in graph theory hasits origin in the four color problem [1]. A graph $G$ is a mathematical structure consisting of two sets $V(G)$ (vertices of $G$) and $E(G)$ (edges of $G$). Proper color- ing of a graph is an assignment of colors either to the vertices of the graphs, or to the edges, in such a way that adjacent vertices / edges are colored differently. Vertex coloring is a hard combinatorial optimization problem.

We apply several operations which act on graphs to give different graphs. In addition to apply graph operations, we color vertices of these obtained graphs properly. Also we developed ColorG package to color the vertices and edges of graphs and to find the Eulerian and the Hamiltonian cycles with web*Mathematica*. Many of these graphs are truly beautiful when drawn properly, and they provide a wide range of structures to manipulate and study.

Before concluding this introduction, we recall some basic definitions.

A complete graph is a simple graph such that every pair of vertices is joined by an edge. A nontrivial closed path is called a cycle. A graph which is obtained by joining a new vertex to every vertices of a cycle is called a wheel. A connected acyclic graph is called a tree [4].

## II. Graph Coloring with Web Mathematica

One of the most exciting new technologies for dynamic mathematics on the World Wide Web is a web*Mathematica*. This new technology developed by Wol- fram research enables instructors to create web sites that allows users to computeand visualize results directly from a web browser. web*Mathematica* is based on a standard java technology called servlets. It allows a site to deliver HTML pages that are enhanced by the addition of *Mathematica* commands [5]. When a request is made for one of these pages the *Mathematica* commands are evaluated and the computed result is placed in the page. People who access web*Mathematica* sites do not have to know how to use Mathematica [11].

In this section, we give applications of ColorG package to color the vertices and the edges of the graphs with web*Mathematica*

### 2.1 Vertex Coloring

The most applications involving vertex coloring are concerned with determining the minimum number of colors required under the condition that the end points of an edge cannot have the same color. A proper vertex coloring of a graph is an assignment from its vertex set to a color set that the end points of each edge are assigned two different colors. The chromatic number of a graph G, denoted by $\chi(G)$, is the minimum number of different colors required for a proper ver- tex coloring of G. Applications of vertex coloring include scheduling, assignment of radio frequencies, separating combustible chemical combinations, and com- puter optimization. We use some commands in the Combinatorica package with Mathematica to color the vertices of graphs and to give web-based examples with webMathematica as in The Fig. 1 [9].

**Fig. 2.** Edge Coloring of a Graph with webMathematica

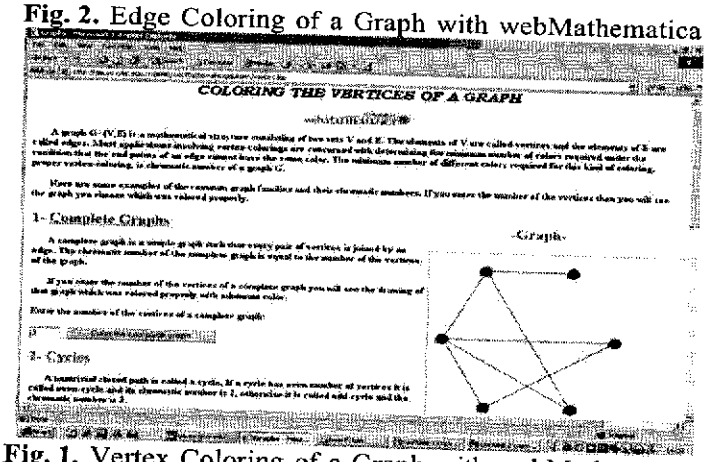



**Fig. 1.** Vertex Coloring of a Graph with webMathematica

If the users press the CompleteGraph, Cycle, Wheel, Star, RandomTree, andany graph and enter the number of vertices they can get the vertex-colored graph.

### 2.2 Edge Coloring

Edge coloring is an optimization problem: An edge-coloring of a graph G is an assignment of colors to the edges of *G* such that edges with a common endpoint

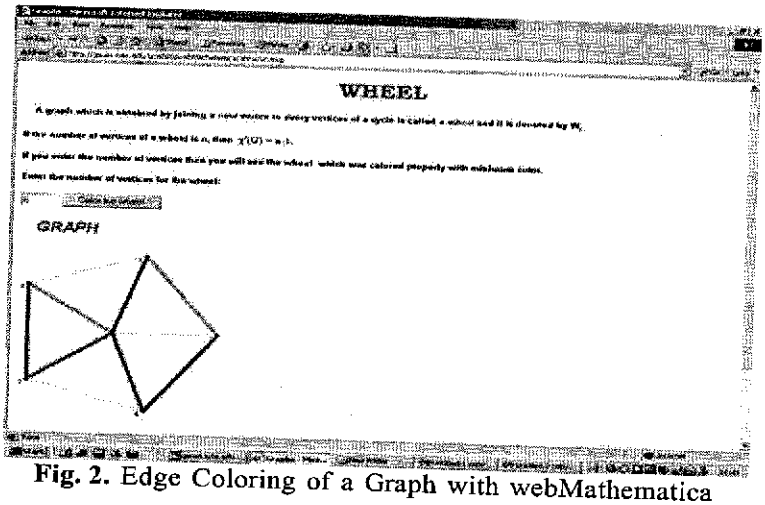



**Fig. 2.** Edge Coloring of a Graph with webMathematica

have different colors. Let $\chi(G)$ denote the chromatic index of $G$, that is the minimum number of colors necessary to color the edges of $G$. Vizing [10] proved that $\chi(G)$ is either $\Delta(G)$ or $\Delta(G) + 1$ for each graph $G$, where $\Delta(G)$ denotes the maximum degree of a vertex in $G$. Then the graph $G$ belongs to one of two classes; either to class 1 or to class 2. This classification problem is NP-complete, and this implies that there are no polynomial-time algorithms for this problem. We use some commands in the ColorG package with Mathematica to color the edges of graphs and to give web-based examples with webMathematica as in The Fig. 2 [9].

If the user press the Wheel button and enter the number of vertices, he/she can get the edge-colored graph.

## III. Cycle Structure in Graphs with WebMathematica

A cycle in a graph is a simple closed path. We will represent a cycle in $G$ as a list of vertices $C = v_1, v_2, \ldots, v_l$ such that there is an edge of $G$ from each vertex to the next in $G$.

### 3.1 Eulerian Cycle

Euler initiated the study of graph theory in 1736 with the famous Seven Bridges of Königsberg problem. The town of Königsberg straddled the Pregel River with a total of seven bridges connecting the two shores and two

islands. The townsfolkwere interested in crossing every bridge exactly once and returning to the startingpoint. An Eulerian cycle is a complete tour of all the edges of a graph. The term circuit is often used instead of cycle, since each vertex can be visited more than once.
We use ColorG package with Mathematica to find the Eulerian cycle and to give web-based examples with webMathematica. If the number of the vertices is entered, it is possible to see the Eulerian cycle in that graph if there exists.

### 3.2 Hamiltonian Cycle

A Hamiltonian cycle of a graph $G$ is a cycle which visits every vertex in $G$ ex- actly once, as opposed to an Eulerian cycle which visits each edge exactly once. A Hamiltonian path is like a Hamiltonian cycle, except that it is a path. Theproblem of computing a Hamiltonian cycle or a Hamiltonian path is fundamen-

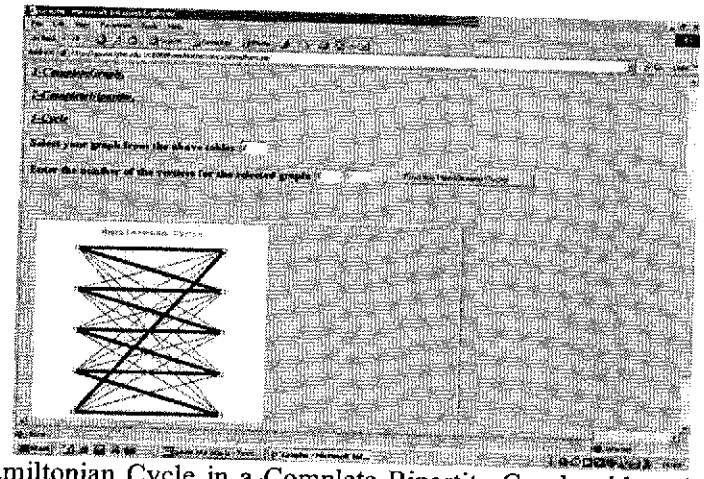


**Fig. 3.** A Hamiltonian Cycle in a Complete Bipartite Graph with webMathematica

tally different from the problem of computing an Eulerian cycle, because testing whether a graph is Hamiltonian is NP-complete.
We use ColorG package with Mathematica to find the Hamiltonian cycle andto give web-based examples with web*Mathematica*. If the number of the vertices is entered, it is possible to see the Hamiltonian cycle in that graph if there exists. The Fig. 4. shows the Hamiltonian Cycle for the Complete Bipartite Graph.

### 3.3 An Application

Some scheduling problems induce a graph coloring, i.e., an assignment of positiveintegers (colors) to vertices of a graph. We discuss a simple example for coloring the vertices of a graph with a small number $k$ of colors and present computational results for calculating the chromatic number, i.e., the minimal possible value of such a $k$.
Draw up an examination schedule involving the minimum number of days for the following problem.
**Example:** Set of students: $S_1$, $S_2$, $S_3$, $S_4$, $S_5$, $S_6$, $S_7$, $S_8$, $S_9$ Examination subjects for each group: {algebra, real analysis, and topology}, {algebra, operations research, and complex analysis}, {real analysis, functional analysis, and topology}, {algebra, graph theory, and combinatorics}, {combinatorics, topology, and functional analysis}, {operations research, graph theory, and coding theory}, {operations research, graph theory, and number theory}, {algebra, number theory, and coding theory}, {algebra, operations research, and real analysis}.
Let $S$ be a set of students, $P$ = 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 be the set of examinations respectively algebra,real analysis, topology, operational research, complex analysis, functional analysis, graph theory, combinatorics, coding theory, and number theory. $S(p)$ be the set of students who will take the examination $p \in P$. Form a graph $G = G(P, E)$, where $a, b \in P$ are adjacent if and only if $S(a)$ $S(b) = \emptyset$. Then each proper vertex coloring of $G$ yieldsan examination schedule with the vertices in any color class representing the schedule on a particular day. Thus $\chi(G)$ gives the minimum number of days required for the examination schedule. The Mathematica commands for this solution are as follows:

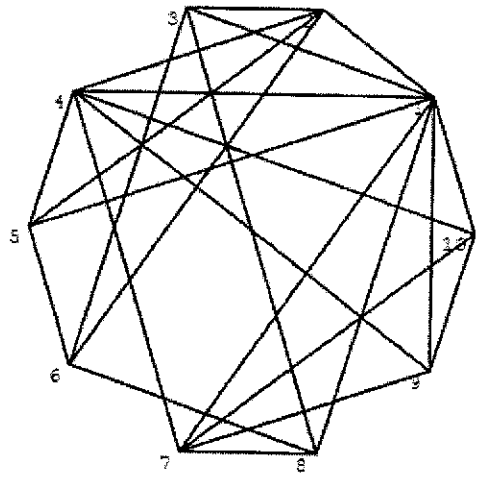



**Fig. 4.** The colored graph of the example

5 days are required and you can see below the lessons in the same parenthesiswhich are on the same day

$$\{\{1, 6\}, \{2, 8, 9\}, \{3, 4\}, \{5, 7\}, \{10\}\}$$

It was very exciting to take 100-year old ideas, simple as they are, and im- plement them in *Mathematica* and web*Mathematica* for anybody. But, there is more work to be done, both of a theoretical and practical nature. If we consider a coloring problem posed as a two-person game, with one person (Alice) trying to color the graph, and the other (Bob) trying to prevent this from happen- ing. Alice and Bob alternate turns, with Alice having the first move. A move consisting of selecting an uncolored vertex $x$ and assigning it a color from the color set $X$ distinct from the colors assigned previously to neighbors of $x$. Ifafter $n = V(G)$ moves, the graph $G$ is colored, Alice is the winner. Bob winsif an impass is reached before all vertices in the graph are colored. The game chromatic number of a graph $G$, denoted by $\chi_g(G)$, is the least cardinality of a color set $X$ for which Alice has a winning strategy [3]. We believe that it is possible to play this game online with web*Mathematica*. Of course, this leads to the rich area of game coloring and the many difficult and intriguing questions there.

## References

[1]. Appel,K.I., Haken,W., and Koch,J.: Every planar map is four colorable I: Dischar-ing, Illinois J.Math, (1977)
[2]. Birkhoff,G.D. and Lewis,D.C.: Chromatic polinomials, Trans.Amer.Math.Soc.,60:355-451, (1946)
[3]. Bodlaender, H.L.: On the complexity of some coloring games, Lecture Notes inComputer Science, (1991)
[4]. Jonathan,G and Jay,Y.: Graph Theory and Its Applications,CRC Press, (1999)
[5]. Pemmaraju,S. and Skiena,S.: Computational Discrete Mathematics, CambrigdeUniv. Press, (2003)
[6]. Skiena,S.: Implementing Discrete Mathematics-Combinatorics and Graph Theorywith Mathematica, Addison-Wesley Publishing Company,(1990)
[7]. Soaty,T.L., and Kainen,P.C.: The four color problem, Dover, New York, (1986)
[8]. Thulasiraman, K. and Swamy,M.N.S.: Graphs: Theory and Algorithms, John Wileyand Sons, Inc., (1992)
[9]. Ufuktepe, U., Bacak, G., and Beseri, T. : Graph Coloring with webMathematica, Lecture Notes in Computer Science, Springer-Verlag, (2003)
[10]. Vizing, V.G.: On an estimate of the chromatic class of a p-graph, Discret. Analiz,(1964)
[11]. Wickham,T.:webMathematica A User Guide,Wolfram Research, Inc., (2002)
[12]. Wolfram,S.:The Mathematica Book, Cambrigde Univ. Press, (1996)