Iraqi Journal for Computer Science and Mathematics

Journal Homepage: http://journal.esj.edu.iq/index.php/IJCM

e-ISSN: 2788-7421 p-ISSN: 2958-0544

FCNN Model for Diagnosis and Analysis of Symmetric Key Cryptosystem

Ali H. Alwan[1,*], Ali. H. Kashmar[1]

[1]Department of Computer Sciences, College of Science, University of Baghdad, Baghdad, Iraq

*Corresponding Author: Ali H. Alwan

DOI: https://doi.org/10.52866/ijcsm.2023.01.01.006

Received August 2022; Accepted October 2022; Available online November 2022

ABSTRACT: An important part of a cryptosystem is a cryptographic algorithm, which protects unauthorized attackers from obtaining private and sensitive data. This study is a research project on identifying cryptographic algorithms using deep learning techniques and categorizing cryptographic algorithms based on feature extraction. The research involves employing block cipher modes called electronic codebook with the encryption algorithms Blowfish and advanced encryption standard (AES), where the data will be encrypted using the same key and a different key. The model has been developed by changing the structure and parameters of the proposed model and the training rate of the data. This model will build several dense FCNN of n layers on regular fully connected neural networks. Its construction will consist of five hidden layers, with each layer consisting of 128 neurons and hidden layers activation Relu except for the output layer, which consists of two classifiers and the SoftMax activation function. FCNN is better able to classify big data. It is also more efficient in use, reducing complexity, with the ability to store training data.

First, the fully connected neural network (FCNN) model was used to evaluate the categorization of the models. Then, all models, even the encryption forms, were evaluated utilizing true positive measurements for satisfactory classification of the identified encryption method and false positive measurements for incorrect classification. The effectiveness of the model was then calculated using the precision value, recall, loss, accuracy range, and F1-Score metrics using a confusion matrix. The FCNN model parameters will be changed to more effectively identify the encryption algorithm. In the proposed method, when using the same key, the accuracy was 81%, and when using a different key, the accuracy was 49%. The FCNN model's adjusted weights and learning will be based on large data to define and assess encryption algorithms more effectively and efficiently.

Keywords: FCNN Models; Diagnosis; Analysis; Symmetric Key; identification; Cryptography Algorithm.

1. INTRODUCTION

Cryptanalysis is a branch of cryptology that focuses on retrieving plaintext from ciphertext with little or no knowledge of the secret key. A cryptanalyst's attempt to decipher the ciphertext is based on the premise that a cryptosystem's secrecy is entirely dependent on the key, and the cryptanalyst is fully aware of the encryption process. However, this is not always the case. The encrypted text is the only information offered to the cryptanalyst [1].

Determining the encryption algorithm only from the ciphertext is thought to be challenging. Very few studies in this sector have been made involving block ciphers or symmetric key ciphers [1]. Cryptoanalysis is used to determine the encryption algorithm from the encrypted text sent by the intruders. Telling which encryption algorithm is being used based just on the ciphertext is impossible. When attempting to decipher the encryption algorithm from the ciphertext, methods based on deep learning and statistics are typically used. The work of identifying the encryption technique is performed using deep learning approaches [2, 3].

A cryptanalyst frequently has no idea which cryptographic technique is being used while collecting particular cipher messages. The first thing a cryptanalyst must do in this case is determine the cryptographic algorithm of the cipher text. The cipher text can subsequently be decoded by the cryptanalyst using a variety of approaches such as brute force, dictionary, rainbow table, and mathematics. Figure 1 depicts the operation of an encryption algorithm identification system based solely on ciphertext knowledge [4].

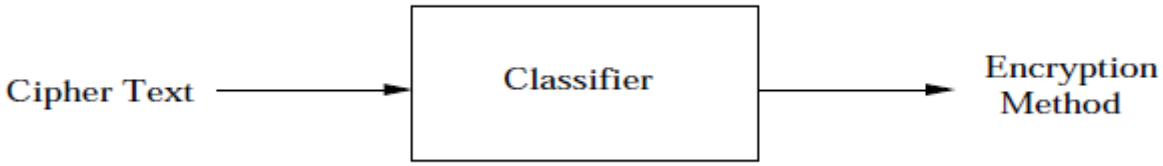

FIGURE 1. Encryption algorithms identification system using just cipher text knowledge [5].

In this study, the encryption algorithms advanced encryption standard (AES) and Blowfish are explored together with tow cryptographic algorithm patterns with mode (electronic codebook, ECB) running. Other modes include cipher feedback mode, output feedback mode (OFB), and counter mode. Our detection methodology is also built on fully connected neural network (FCNN) technology, a type of deep learning-based approach. The form must first be used to extract the features of each ciphertext. After entering the ciphertext into the identification form, the proper encryption algorithm can be established.

2. BACKGROUND

2.1 DEEP LEARNING

Deep learning has grown in popularity in scientific computing, and firms that deal with complex problems commonly use its methodologies. All deep learning methods use various types of neural networks to do certain tasks. Deep learning uses artificial neural networks to do difficult calculations on massive amounts of data. Deep learning is a type of artificial intelligence based on the structure and operation of the human brain. Deep learning approaches instruct computers through the use of examples and is widely employed in healthcare, e-commerce, media, and entertainment. A neural network, which is arranged similarly to the human brain, is made up of artificial neurons, also known as nodes. As indicated in Figure 2, the three layers that these nodes are piled on top of one another are the input layer, the hidden layer(s), and the output layer [6].

Each node receives information in the form of data inputs. The inputs are computed, multiplied by random weights, and a bias is applied. Nonlinear functions, also known as activation functions, are utilized to determine which neurons to activate. Although self-learning representations distinguish deep learning algorithms, they also rely on artificial neural networks (ANNs) that simulate how the brain processes information. Throughout the training phase, algorithms utilize unknown aspects in the input distribution to extract features, classify objects, and find meaningful data patterns. This occurs on multiple levels, with the algorithms used to generate the models, much like teaching machines to learn for themselves. Deep learning models employ a number of algorithms. No network appears to be faultless, and some algorithms appear to be better suited to specific jobs. To make the best decisions, a complete understanding of all main algorithms, such as convolutional neural networks (CNNs), recurrent neural networks, and FCNNs, is useful [6].

2.2 SYMMETRIC KEY ALGORITHMS USED WITH IN OUR EXPERIMENTS

2.2.1. BLOWFISH

Blowfish was invented by Bruce Schneier. Although it uses symmetric keys and is similar to DES, the key size can range from 32 to 448 bits. It is a block cipher of 64 bits. The two components of the Blowfish algorithm are key expansion and data encryption. Key expansion, in which the key is divided into numerous subkeys, precedes data encryption. A Feistel function is used in the data encryption process, and it is repeated 16 times, or 16 rounds, of the method. A key-dependent permutation and a key-and-data-dependent replacement are both used in each cycle. The S- array uses the four bytes that the Feistel function divides from the 32-bit input as its indices. The lookup results are then added, and their XOR operation is then carried out to produce the output [8].

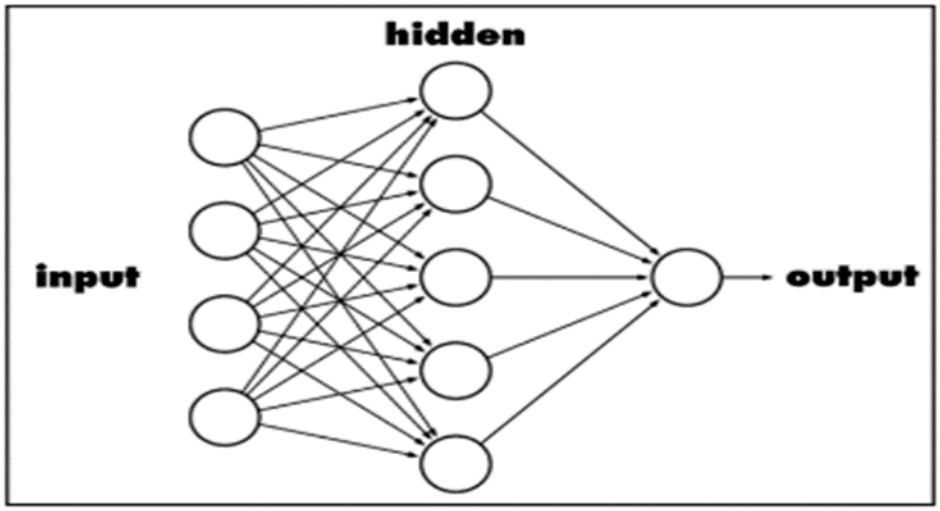

FIGURE 2. Structure diagram neural network [7].

2.2.2. ADVANCED ENCRYPTION STANDARD

NIST introduced AES or Rijndael in 2001. It is a symmetric key algorithm that runs for 10, 12, and 14 cycles with keys of 128, 192, and 256 bits, respectively. AES uses permutations and combinations in each round. A 4*4 matrix, commonly referred to as a state matrix, is used by AES. To perform rounds, AES divides the data into four blocks or an array of bytes that create 4*4 matrices. AES has three rounds: an initial round, a final round, and a key expansion. The same procedures must be carried out on the state matrix during every AES cycle except from the final one, as follows [8]:

1- Substitute bytes: These comprise processes that are built on specially created substitution boxes. This operation's primary goal is to stop numerous attacks, including mathematical attacks, differential cryptanalysis, and linear cryptanalysis.

2- Shift rows: This fundamental linear procedure is carried out on state matrices the intention of causing diffusion.

3- Mix columns: Similar to the shift row operation, this is likewise a fundamental operation. Matrix multiplication is involved.

4- Add round key: This simple procedure conducts an exclusive-or operation between the state matrix and the key of that round. The outcome is applied in the following round. This operation's goal is to cause confusion [8].

3. RELATED WORK

Nagireddy, S. **In (2019)** [9]. This study proposed identification of the encryption techniques of block ciphers is a type of cryptanalysis attack. This is the main attack used to evaluate the security of block ciphers. This task has two issues: (1) encryption model is identified and (2) determining the encryption methods used by block ciphers is a pattern classification issue. Early methods for identifying patterns in block ciphers included the histogram matching approach, support vector machines, and Gaussian mixture models. to create a variety of attacking methods, including ciphertext- only assault, known-plaintext attack, and side-channel attack.

Hu, X., & Zhao, Y. **In (2019, February)** [10]. The results demonstrate that the generated features can successfully extract from the contents of ciphertext data. This research proposed a novel method by constructing a feature based on ciphertext recombination and location-detection. Eight algorithms can correctly categorize objects in ECB mode with an accuracy of more than 87%. In the CBC mode, it is also capable of classifying more accurately than randomly. Eight block ciphers of the electronic code book mode (ECB) and eight block ciphers of the cipher block chaining mode were classified using a random forest classifier (CBC).

Kopal, N. **In (2020, May)** [11]. The cryptanalysis in this research is the development of an artificial neural network that is able to recognize between five classical ciphers: simple monoalphabetic substitution, Vigenère, Playfair, Hill, and transposition. The network is based on Google's TensorFlow library as well as Keras. This paper presented the current progress in the research on using such ANNs for diagnosis the cryptography type. ANNs can categorize about 90% of the ciphertexts correctly. using an artificial neural network that provides 54% computation time for categories of cipher kinds.

Yang, W., & Park, Y. **In (2021, Jan)** [12]. This study presents a new technique that can identify symmetric key algorithms using CNN for traces extracted from Intel processor trace (IPT). The IPT interrupts the running of software and IPT is used to extract the trace that has been symmetric key ciphered first. It is then transformed into an image and fed into CNN. Two different datasets were used for the experiment. The training results from the first dataset, which contained traces extracted using various symmetric key algorithms, were categorized into nine classes with 100% accuracy. The second dataset contained traces for each type of symmetric key algorithm, along with the block cipher modes and various bit sizes of the security keys. The accuracy of the classification of the training results into 36 classes was 70.55%. This study used a CNN to determine the number of key bits and the block-cipher modes in addition to the types of encryption algorithms that have been previously identified by studies.

4. PROPOSED MODEL

In this study, a deep learning strategy is suggested for determining the encryption algorithm from the cipher text, using FCNNs. In FCNN, nodes are usually referred to as "neurons." An FCNN is made up of several fully linked layers. A fully connected (FC) layer is a function that ranges from R m to R n. Each output dimension is influenced by each input dimension [5]. After the CONV layers, a CNN may have FC layers. These layers, like the traditional multi-layer perceptron, perform vector transformation and serve as a classifier [6]. FCNNs are efficient in reducing complexity, as well as working on large data and storing training data, so they are more effective [13].

5. EXPERIMENTAL PROCESS

5.1 DATA SET

Using cryptographic algorithms from the Kaggle website, which contains plain text and ciphertext, this research paper's ciphertext will be constructed from plain text data using an unnamed dataset [14]. A single key or several keys is used to encrypt data. The dataset's original texts will first be converted into ciphertexts using cryptographic techniques. The data will then be encrypted and kept in a dictionary, with training data discounted by 80% and testing data discounted by 20%. The unlabeled dataset will be used to evaluate the model using a confusion matrix to identify the coding algorithm after it has been trained and executed. For this investigation, 1,000 samples were randomly selected from a sample of 45,133 samples.

5.2 SYSTEM MODEL

In this study, the FCNN deep learning method was used. The experiment started with preparing encrypted data for the FCNN model's training, cross validation, and testing. In the end, a set of accuracy measurements were used to evaluate the forms. Utilizing the SoftMax activation function in the FCNN model. A few standard model parameters can be altered when using FCNN to solve classification problems. This includes changing the weights and amount of layers. The learning rate might be adjusted during the training period to maximize form development. As part of the experiment's design and model optimization, these were changed. However, because one of the research hypotheses is to look at how the depth of the deep learning method affects this classification problem, the depth of the forms was modified, or, more precisely, raised. The key property of FCNNs is deep architecture. Other recent experiments demonstrate that the algorithms get more adept at seeing minute patterns in encrypted data as the number of layers increases. A total of 1000 samples and 100 epochs were used to train each model. The lengthy training duration and the scarce research resources were the main causes of the brief training cycle. By the middle of the training stage, the validation loss and accuracy have reached saturation, but the next training visualization group shows that the training epochs were sufficient. Owing to the pre-training model design, the form responded quickly to the training dataset. The data are preprocessed before being used to train the FCNN network, as seen in Fig. 4. After being recovered by FCNN, SoftMax classifies the outputs of FCNN as well as the feature data.

The most essential feature of FCNN is its deep architecture. According to other recent studies, the deeper the layers, the better the algorithms are at detecting small patterns in encrypted data. The general layout of the suggested approach for the analysis and identification of symmetric key algorithms is depicted flowcharts in **algorithm 1**. The following parts provide specifics into the suggested method.

6. EVALUATION

The trained models were evaluated or tested using a test data set that used the same set of cryptographic algorithms that will use the same key or a different key at random. It will also be used with block cipher modes of operation used

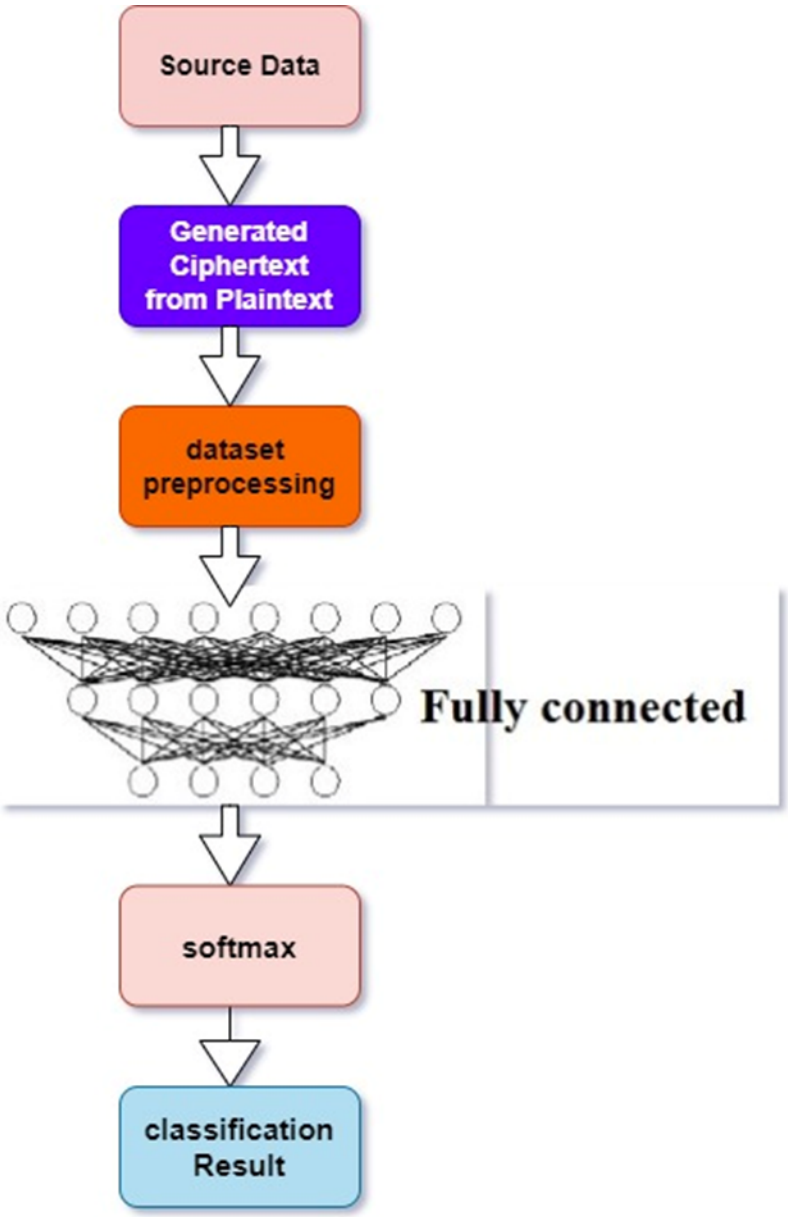

FIGURE 3. **Structurediagram of FCNNs.**

ALGORITHM 1. **DIAGNOSIS AND ANALYSIS FOR SYMMETRIC KEY CRYPTOSYSTEM**

Input	DNN algorithm, encryption algorithm 1 and encryption algorithm 2
Output	Metrices and confusion matrix.
Step 1	FCNN algorithm type, encryption algorithm 1 and encryption algorithm 2, get data from dataset.
Step 2	Generate random plaintext (paragraph) 1000 to 3000.
Step 3	Generate cipher from plaintext for two crypto algorithms in same time by using same key, different key.
Step 4	Get cipher text generated (En1+En2), num of trains (epoch) 50, number of steps in each train (iteration) 100,
Step 5	Get cipher from dataset for test
Step 6	Evaluate trained test results
Step 7	Optimized Results
Step 8	Show final results
Step 9	End

in cryptography, including the ECB mode. The evaluation was carried out using two different encryption systems. The comparison list used in the test is shown below.

Table 1. The Comparison List Used in The Test.

Electronic codebook (ECB) mode
Blowfish_or_AES_Cipher Same key
Blowfish_or_AES_Cipher_ Different Keys

The FCNN model was used to perform the classification evaluation of the base models first. Then, all forms, including the forms, were evaluated using true positive measurements for valid classification of the encryption identification scheme and false positive measurements for incorrect categories. The precision, recall, loss, accuracy, and F1-score (abbreviated as F1) the following formulas were then used to calculate the measurements [15].

$$\text{PRECISION} = \frac{TP}{TP+FP} \tag{1}$$

$$\text{RECALL} = \frac{TP}{TP+FN} \tag{2}$$

$$ACCU\ RACY = \frac{TP+TN}{TP+TN+FP+FN} \tag{3}$$

$$F1 = 2 * \frac{\text{PRECISION} * \text{RECALL}}{\text{PRECISION} + \text{RECALL}} \tag{4}$$

The classification tests of two encryption techniques were used in this study and is easily extensible to perform multiclass categorization. Following the completion of the experiments, the categorization results were analyzed to assess the performance of the trained forms. The first set of analyses is to evaluate all forms' categorization. The evaluation of the confusion matrix of the basic models used in this research, as well as the accuracy and loss, to identify and analyze symmetric key algorithms using FCNN are shown in the tables below.

This study validated the possibility of identifying symmetric key algorithms by using an FCNN on the Blowfish vs. AES model with the same key and ECB mode block cipher. The accuracy in identifying the types of symmetric key algorithms using the same key with ECB mode was 81%, and the loss was 40%, as illustrated in Table 2 and Figure 6.

Table 2. Blowfish_or_AES_Cipher_Same key_ ECB Mode

Accuracy	Loss	Precision value	Recall	F1 score	Support
0.81	0.40	0.81	0.81	0.81	1000

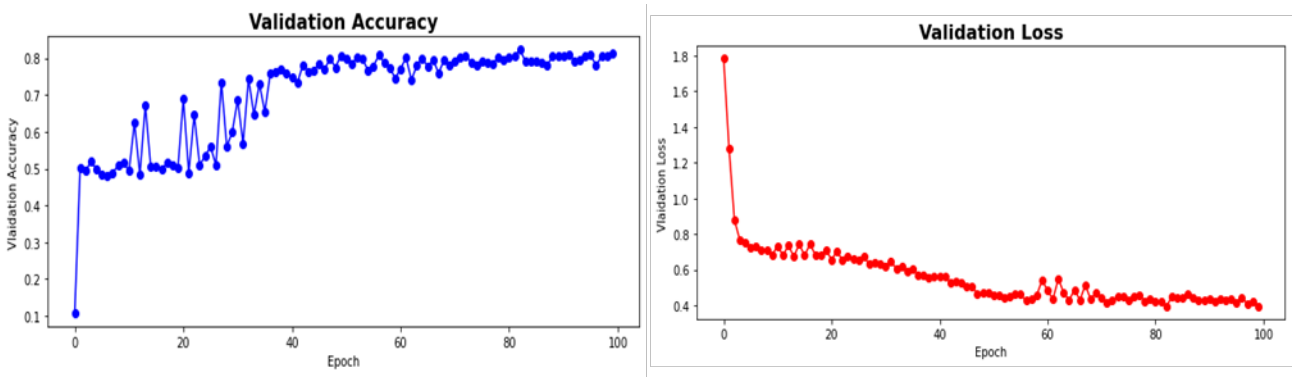

FIGURE 4. Accuracy and Loss Blowfish_or_AES_Cipher_Same key_ CBC Mode.

Figure 6 illustrates this. The confusion matrix is a (N x N) matrix that is used to assess the performance of a categorization form, where (N) is the number of target classes. An individual could determine the accuracy of the model by observing the diagonal values for measuring the number of correct categorizations by visualizing the confusion matrix. The model has been proven to be successful.

Table 2 demonstrates the possibility of detecting symmetric key algorithms by using an FCNN on the Blowfish vs. AES model, which uses the same key as the ECB mode block cipher. As shown in Figure 5, the performance was 81%

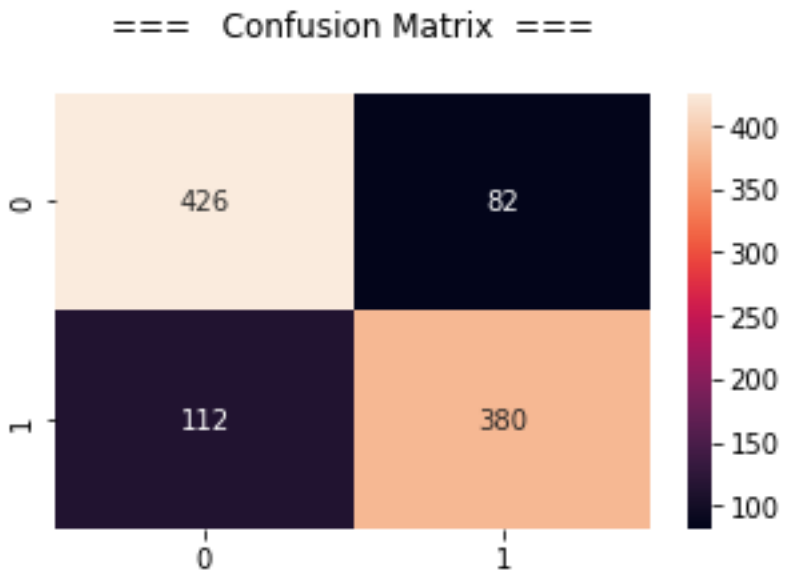

FIGURE 5. Confusion Matrix Blowfish_or_AES_Cipher_Same key_ ECB Mode.

accurate in identifying the types of symmetric key algorithms using the same key with ECB mode and indicated 40% loss. Furthermore, as shown in Figure 6, a confusion matrix was used to determine whether the model was successful.

The ability to detect symmetric-key algorithms is shown in Table 2. Different keys are used with ECB mode block ciphers by Blowfish vs. AES. Accuracy in classifying the various symmetric key algorithms using a different key and ECB mode was 49%, and a loss of 70% was observed, as shown in Figure 7.

Table 3. Blowfish_or_AES_Cipher_different key _ ECB Mode

Accuracy	Loss	Precision value	Recall	F1 score	Support
0.49	0.70	0.50	0.50	0.44	1000

Figure 7 shows that although the accuracy range is 49% and the loss is 70%, using a different key and the suggested model to assess the security algorithm demonstrates the algorithm's strength in terms of security. Additionally, the confusion matrix, which used 1000 samples and displayed the number of recognized and unrecognized samples, as shown in Figure. 8, helped to clarify the situation.

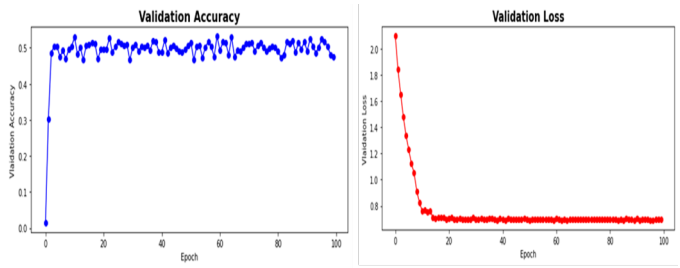

FIGURE 6. Accuracy and Loss AES_or_DES3_Cipher_ different key _ CBC Mode.

7. COMPARING OUTCOMES

The average accuracy, loss, precession value, recall, and F-score were used to compare the results of the models in Tables 2 and 3, displaying the highest potential value reached and the worst outcome of the models when employing the suggested technique. The comparison will show the evaluation of the encryption algorithm and its security and confidentiality, as shown in the scheme below (figure 9).

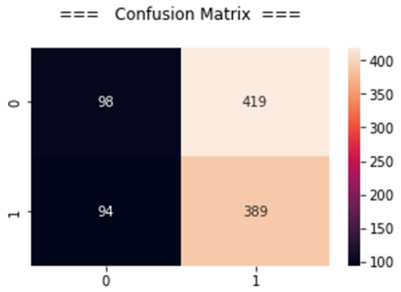

FIGURE 7. Confusion Matrix AES_or_3DES_Cipher_ Different key _ CBC Mode.

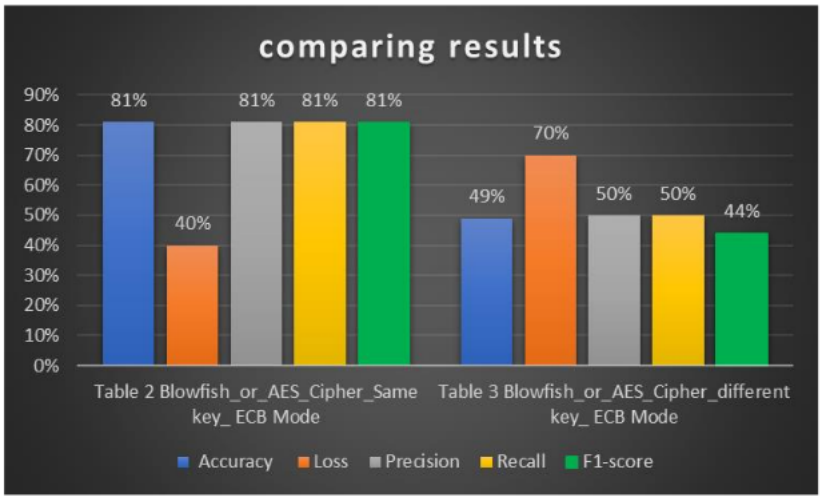

FIGURE 8. comparing results

The suggested technique, which uses the same key and a different key in encryption with ECB mode, is superior to the source [16], which also uses the same key in encryption with the ECB mode that uses rotation forest, according to a performance comparison shown in Table 4.

Table 4. Comparison of classification accuracy of block ciphers same KEY IN ECB mode.

Mode	The proposed method Accuracy same key	Source [16] Accuracy Rotation Forest same key
ECB	81%	54%

8. CONCLUSIONS

In a variety of industries, categorization issues are addressed using artificial intelligence techniques like deep learning. The deep learning algorithm with the best classification performance is specifically identified as FCNN. This paper introduced a novel approach for categorizing symmetric key methods using a deep learning model to identify an encryption technique from a ciphertext (FCNN). The FCNN model, a multi-layered feedforward neural network model, is one of the most important deep learning models. FCNN uses the spatial concept to extract characteristics. The results demonstrated that an encryption algorithm that utilizes the same key performs better, with 81% accuracy, than an algorithm that uses a different key, thereby providing a means of gauging block cipher strength for the study. Our research has demonstrated that, unlike models that use a separate key that is more secret, encryption approaches that utilize the same key for encryption do not provide good message confidentiality. with epoch 100, and 1000 samples as the sample size. The characteristics of cipher algorithms also formed the basis for this study, enabling the recognition, classification, and examination of the symmetric master cipher system's algorithm type.

ACKNOWLEDGEMENT

The first author would like to thank the reviewers for providing useful suggestions, allowing for the improved presentation of this paper.

CONFLICTS OF INTEREST

The authors declare no conflict of interest.

REFERENCES

[1] K. V. Pradeepthi, V. Tiwari, and A. Saxena, "Machine learning approach for analysing encrypted data," *Tenth International Conference on Advanced Computing (ICoAC)*, 2018.
[2] S. Swapna, "Block cipher identification using support vector classification and regression," *Journal of Discrete Mathematical Sciences and Cryptography*, vol. 13, pp. 305–318, 2010.
[3] A. Dileep, C. C. Dinesh, and Sekhar, "Identification of block ciphers using support vector machines," *The 2006 IEEE International Joint Conference on Neural Network Proceedings. IEEE*, 2006.
[4] C. Tan and Q. Ji, "An approach to identifying cryptographic algorithm from ciphertext," *8th IEEE International Conference on Communication Software and Networks (ICCSN)*, 2016.
[5] S. Nagireddy, 2008.
[6] A. Biswal, 2021.
[7] G. Ciaburro and B. Venkateswaran, 2017.
[8] M. Sohal and S. Sharma, "BDNA-A DNA inspired symmetric key cryptographic technique to secure cloud computing," *Journal of King Saud University-Computer and Information Sciences*, vol. 34, no. 1, pp. 1417–1425, 2022.
[9] S. Nagireddy, "A pattern recognition approach to block cipher identification," 2019. LAP LAMBERT Academic Publishing.
[10] X. Hu and Y. Zhao, "Block ciphers classification based on random forest," *Journal of Physics: Conference Series*, vol. 1168, no. 3, 2019.
[11] N. Kopal, "Of ciphers and neurons-detecting the type of ciphers using artificial neural networks," *Proceedings of the 3rd International Conference on Historical Cryptology HistoCrypt*, 2020.
[12] W. Yang and Y. Park, 2021.
[13] B. Ramsundar and R. B. Zadeh, 2018.
[14] D. K. Website.
[15] X. Zhou, J. Feng, and Y. Li, "Non-intrusive load decomposition based on CNN-LSTM hybrid deep learning model," *Energy Reports*, vol. 7, pp. 5762–5771, 2021.
[16] T. Kavitha, "Classification of encryption algorithms based on ciphertext using pattern recognition techniques," in *International conference on Computer Networks, Big data and IoT*, Springer, 2018.