International Journal of Computer Information Systems and Industrial Management Applications
ISSN 2150-7988 Volume 18(17s) 2026 pp. 1360-1373
https://doi.org/10.70917/ijcisim-2026-4832
Article

Cybersecurity Framework using Deep learning to Detection of Malicious 2D Barcode Attacks

Pranav Pradeep Menon[1], Devraj Vishnu[1], Narottam Das Patel[1], Praveen Lalwani[1]

[1] School of Computer Science and Artificial Intelligence, VIT Bhopal University, Bhopal, India
Email: pranav.pradeep2021@vitbhopal.ac.in
Email: devrajvishnu@vitbhopal.ac.in
Email: narottamdaspatel@vitbhopal.ac.in
Email: praveen.lalwani@vitbhopal.ac.in

Abstract: Two-dimensional (2D) barcodes are being used in the development of many modern digital services, such as payment systems, authentication systems, and information access systems. However, the use of the barcode has been exploited by attackers, who insert malicious payloads in the barcode while keeping the structure of the barcode unchanged. This article proposes the application of deep learning model for detecting visually manipulated barcodes directly from the image data. Two models, namely a Convolutional Neural Network (CNN) model and a hybrid CNN and Long Short-Term Memory (LSTM) model, are considered for the purpose of detecting visually manipulated barcodes. A dataset of 80,160 images of barcodes is created, which includes 16,032 benign barcodes and four types of attack scenarios: encoded manipulation, payload injection, structural rotation, and adversarial noise. The results obtained from the experiment show that the CNN model has achieved an accuracy of 99.97%, and the CNN-LSTM model has achieved an accuracy of 99.98%.
Keywords: Two-dimensional (2D) barcodes, Convolutional Neural Network, Long Short-Term Memory, Malicious

1. INTRODUCTION

Two-dimensional (2D) barcodes, like QR codes and similar visual mechanisms of encoding, are now an essential part of the digital world. These are commonly employed in various applications such as mobile payment systems, product verification, tracking, ticketing systems, and accessing online services quickly. These are popular because of their ability to store a considerable amount of data and are easily readable using mobile devices and imaging technology at minimal cost [6]. In the current scenario of the increasing popularity of contactless technology and smartphone-based applications, barcode technology has emerged as a vital link between the real and virtual worlds.

However, there are some security issues that can be raised with the barcode system, despite its flexibility and usability. The barcode system, which is based on the scanning mechanism and the automatic reading of the information embedded in the barcode, can be exploited by attackers to embed malicious information and instructions in the barcode. This is because the barcode scanning system does not perform any visual

verification of the structure of the barcode; rather, it automatically reads the information embedded in the barcode and executes the instructions embedded in it. Therefore, attackers can easily exploit the barcode system by modifying the visual structure of the barcode and embedding malicious instructions in it. Such malicious instructions can be difficult to detect because the barcode is visually correct and can be scanned normally.

Some security issues that can be raised with the barcode system have been identified in recent studies, which include content manipulation, malicious URL embedding, and structural tampering attacks [5].

In conventional security mechanisms, the main emphasis is placed on verifying the accuracy of the decoded information obtained after scanning the barcode. In this regard, conventional security mechanisms are based on

checking the accuracy of the extracted information against known malicious URLs and patterns. Although conventional security mechanisms can identify known malicious activities, they are ineffective in checking the visual integrity of the barcode images. In this regard, sophisticated manipulations of the barcode images can evade conventional security mechanisms, as they are still decodable and can evade conventional security mechanisms as discussed in [5].

In the domain of cyber security, image-based deep learning techniques have also been employed for malware detection by converting malware information into visual forms and analyzing them through CNN-based models [9]. All these techniques clearly indicate the potential benefits associated with visual pattern analysis for malware detection. In addition, deep learning techniques have been employed for barcode recognition and scanning systems by employing CNN-based techniques for enhancing barcode detection and decoding systems [10]. All these techniques clearly indicate the potential benefits associated with deep learning-based image analysis for enhancing barcode security systems.

Though CNNs perform well in feature extraction from spatial information, they might fail to effectively grasp the sequential dependencies often embedded in visual structural patterns. The Long Short-Term Memory (LSTM) networks are a variant of RNNs used for effectively grasping sequential dependencies in data [2]. LSTMs make use of memory cells in addition to gates in order to effectively grasp complex dependencies in data sequences [3]. By effectively combining CNNs for feature extraction with LSTMs for grasping sequential dependencies in data sequences, this approach can effectively grasp both spatial and sequential dependencies in visual structural patterns, thereby effectively detecting subtle structural changes in visual patterns.

Based on the above facts, the framework for the detection of malicious changes in visual structural patterns with the use of deep learning has been proposed in this paper. Two models are considered in the context of evaluating the framework: one CNN model for the purpose of feature extraction with the use of visual structural patterns, and the CNN-LSTM model for the purpose of feature extraction and grasping dependencies with the use of visual structural patterns.

To assess the efficacy of the suggested approach, a synthetic dataset was created based on a publicly accessible barcode image dataset [12]. A total of 16,032 benign images of barcodes were used as a base dataset. Four different types of attacks were created based on each image: encoded manipulation, payload injection, structural rotation, and adversarial noise injection. This resulted in a synthetic dataset consisting of 80,160 images of barcodes, divided into five classes.

Experimental results show that the suggested approach based on deep learning models can effectively classify malicious manipulations of images of barcodes. In the experiment, a standalone CNN model was used, achieving a classification accuracy of 99.97%. In addition, a hybrid model based on a combination of CNN and LSTM models was used, achieving a

classification accuracy of 99.98%. This shows that image-based deep learning models are a promising approach to solving the cybersecurity problem of detecting malicious manipulations of images of barcodes.

2. RELATED WORK

The rapid acceptance and usage of two-dimensional (2D) barcodes in modern digital technologies have encouraged research in the security of barcodes and the detection of malicious content. Barcodes are commonly employed in financial transactions, security systems, and information access control. This has encouraged the use of barcodes in critical systems and has increased the need for the security of barcodebased communication systems. Barcode-based security has therefore become an important research area in recent years as the usage of barcodes in modern systems has increased [6].

The early research in barcode security was based on the analysis and detection of the content received after decoding the barcode. The traditional detection systems commonly employed rule-based filtering and blacklist-based

verification and signature-based detection systems for the detection of malicious content such as URLs, scripts, and text data embedded in the barcodes [5]. The content received after decoding the barcode is analyzed and checked for its compliance with the known malicious content and the detection systems are based on the analysis and detection of the content received after decoding the barcode.

Several studies works have investigated the effectiveness of making use of machine mastering techniques to conquer the barriers of traditional barcode safety features, in particular in identifying malicious content material embedded in barcodeprimarily based assaults. Classical device studying techniques which include guide Vector gadget (SVM), Random woodland classifiers, and kNearest Neighbour (k-NN) are applied to discover suspicious styles within the content of the barcode or URLs of the content [8]. these methods typically rely on manually engineered features derived from textual residences, statistical traits, or behavioural attributes of the decoded statistics. at the same time as such strategies frequently gain improved detection accuracy compared to rule-based filtering strategies, their effectiveness in large part relies upon on the quality and relevance of the handcrafted functions. consequently, their overall performance can also decline when encountering previously unseen or evolving attack strategies [8].

With the fast development of artificial intelligence, deep gaining knowledge of techniques have attracted widespread hobby inside the discipline of cybersecurity studies, particularly in information evaluation based on pics. specific from traditional gadget mastering techniques, deep gaining knowledge of methods can research directly from uncooked records without any guide intervention in feature extraction. among all deep learning strategies, Convolutional Neural Networks (CNNs) have demonstrated amazing overall performance in photograph class and visible pattern popularity because of their ability to research hierarchical features from pictures [4], [11]. In cybersecurity applications, CNN-based techniques have been efficaciously used for malware detection with the aid of converting binary executable documents into grayscale photo representations and extracting discriminative patterns that distinguish malicious software program from benign applications [9]. those findings spotlight the effectiveness of visible pattern evaluation in revealing hidden structural characteristics found in malicious samples.

Motivated by these advancements, recent studies have begun the investigation of the application of image-based deep learning models for barcode analysis and recognition. CNN-based models have been employed for barcode and QR code recognition, where the networks are trained to learn spatial representations of barcode patterns for enhanced accuracy and effectiveness in barcode recognition [10]. These studies demonstrate the feasibility of treating barcode images as visual patterns rather than mere containers of textual data.

If you want to similarly improvise the gaining knowledge of capability of deep gaining knowledge of fashions, hybrid models of CNNs and recurrent neural networks have been proposed for photograph analysis and know-how. lengthy quick time period memory (LSTM) is a form of Recurrent Neural network that is capable of mastering sequential relationships and long-term dependencies in statistics sequences[2]. LSTM is primarily based at the potential of the networks to examine and model sequential relationships and contextual relationships which might be present within the data with regards to diverse functions[3]. Hybrid CNN-LSTM models have been successfully employed for various applications, including malware analysis, network traffic classification, document analysis, and intrusion detection systems, where sequential relationships and contextual relationships need to be learned and modelled in the data.

Despite these advancements, little research has been conducted to specifically identify malicious manipulations in 2D barcode images using deep learning methods. Existing studies are mostly focused on increasing the accuracy of barcode image recognition or exploring the content of the barcode images after they are decoded. In this regard, malicious barcode images with visual manipulations, which are still decodable, can evade current security systems.

Therefore, it is evident that there is a need for research in this area to identify and explore just how effective it is to use deep learning methods which can learn the spatial and structural characteristics of barcode images for security analysis purposes. In this paper, we try to bridge this gap of research by exploring the effectiveness of a standalone Convolutional Neural Network (CNN) and a hybrid CNN-LSTM approach in detecting malicious barcode manipulations.

3. BACKGROUND AND DATASET CONSTRUCTION

Here, we will discuss the structural properties of 2D barcode images, as well as the methodology for building the dataset used in this paper. The dataset was built from a publicly available repository on Kaggle, and synthetic attacks were added to simulate different kinds of attacks on barcodes. Using this method, it is possible to create a balanced dataset for training deep learning models for image recognition tasks.

A. Two-Dimensional Barcode Representation

Define 2D barcodes are capable of storing a larger amount of data compared to traditional 1D barcodes, as they are able to represent information along both horizontal and vertical axes. A standard 2D barcode consists of a combination of black and white modules, which are organized in a grid-like pattern, with the pattern of modules used to represent information.

A critical aspect of 2D barcodes is the incorporation of error correction codes, which enable the barcode to remain "readable" even if a part of the image is damaged. Although this provides a level of practical advantage, this also poses a potential security risk, as the ability of a 2D barcode to remain "readable" even after alterations allows an attacker to modify parts of the barcode structure while still keeping it "scannable."

For example, an adversary could modify pixel patterns or make minor structural changes to the barcode, which would still make it decodable but could alter the data embedded in it. In such cases, it is possible that the security measures in place might fail to identify the manipulation of the barcode, as they are based on the analysis of the data extracted from it. Hence, it is important to analyze the visual structure of the barcode images in order to identify malicious modifications. This is where image-based deep learning can prove to be effective.

B. Creation of the Dataset

The base barcode dataset used in this research was obtained from the Kaggle repository titled “AI Generated Barcode Images & Masks Dataset” [12]. The dataset was downloaded and stored locally in the directory path D:\archive, which contains two main folders: one containing barcode images and another containing segmentation masks.

Only the barcode images were used for the classification task in this study, while the mask images were not utilized. A total of 16,032 barcode images were extracted from the images directory to form the initial dataset of benign barcode samples.

Before training the deep learning models, the images were pre-processed to ensure consistency in input dimensions and pixel values. Each image was converted to grayscale format and resized to 128 × 128 pixels. In addition, pixel intensity values were normalized to the range [0,1] to facilitate stable neural network training and improve convergence during optimization.

C. Attack Simulation and Dataset Expansion

To simulate the attacks of malicious barcode manipulations, four synthetic attack techniques were used to manipulate each benign barcode image. These attack techniques were designed to simulate visual distortions in the barcode images while preserving their general structure and scannability. Four images were created for each benign barcode image, making a total of five classes: one class for the benign images and four classes for the attack images.

After applying the attack techniques to the images, the total dataset was expanded from 16,032 original images to 80,160 images. The five classes in the dataset are:

Benign Barcode – Original image of the barcode.

Encoded Manipulation Attack – In the encoded manipulation attack, 300 random pixel positions in the barcode image have been chosen, and the intensity values have been modified. This type of attack causes distortion in the structure of the barcode.

Payload Injection Attack – In the payload injection attack, the 64x64 region of the barcode image has been replaced with random values ranging from 0.8 to 1. This type of attack simulates the injection of malicious visual patterns in the barcode structure.

Structural Rotation Attack – In the structural rotation attack, the barcode image has been rotated by applying a 5-degree affine transformation around the center point.

Adversarial Noise Attack – In the adversarial noise attack, the image has been modified by adding signed noise with a magnitude factor of 0.03. This type of attack simulates the adversarial attack on machine learning models.

These synthetic attacks enable the generation of diverse malicious barcode patterns and allow the deep learning models to learn structural differences between legitimate and manipulated barcode images.

D. Training and Testing Split

After the dataset was expanded, the entire dataset was shuffled at random and then split into the training set and the testing set using an 80-20 split. Stratified sampling was used in this case so that all classes were adequately represented in the dataset.

From this dataset, sixty four,128 pictures have been used for the model's training, and 16,032 pictures had been set apart for the model's trying out and assessment. at some stage in the version's training, a part of the dataset became set aside as the validation set so that the version's performance become monitored and overfitting became prevented.

This dataset become organized in this kind of way that it's far wellbalanced and various and can be used for the evaluation of the overall performance and effectiveness of the model in detecting any form of malicious changes in the barcodes.

4. PROPOSED FRAMEWORK

The proposed framework is expected to analyze the barcode images and identify whether the image is associated with a benign barcode or one of the malicious manipulation types identified in this paper. Unlike previous approaches, which analyze the content of the barcode after it has been decoded, the proposed approach will utilize direct image analysis of the barcode using deep learning technology, as discussed in [6] and [7]. The workflow of this proposed system is illustrated in Fig. 1, where two different architectures of deep learning are utilized: one based on a Pure Convolutional Neural Network (CNN), and the other based on a hybrid CNN-Long Short-Term Memory (CNN-LSTM).

Prior to education the deep mastering fashions, a preprocessing level is carried out to all barcode pics to standardize the input facts. at some point of this level, each photo is first transformed to grayscale after which resized to a

set measurement of 128 x 128 pixels. Standardizing the picture dimensions ensures uniform enter formatting simply so the neural network can machine the information effectively. furthermore, pixel intensity values are normalized to the range [0,1], which allows decorate numerical balance in some unspecified time in the future of schooling and facilitates faster convergence of the optimization algorithm. Such preprocessing operations are extensively adopted in deep gaining knowledge of–primarily based photo classification pipelines to enhance version training overall performance and reliability [4].

Following preprocessing, the prepared dataset consists of five categories of barcode images: benign barcodes, encoded manipulation attacks, payload injection attacks, structural rotation attacks, and adversarial noise attacks.

These categories represent both legitimate barcode samples and various forms of malicious visual manipulations, as illustrated in Fig. 1.

The first architecture that is being examined in this study is the Convolutional Neural Network (CNN), which is intended to automatically learn spatial features from the images of the barcodes. The convolutional layers in the CNN utilize filters that can be learned and can be slid over the image to automatically learn visual patterns such as edges and shapes that can be seen in the barcode image [1]. In the early layers of the network, basic visual patterns can be observed, and as the layers go deeper, complex patterns can be seen that pertain to the barcode structure itself. Batch normalization is also included in the layers to stabilize the learning process. Max-pooling operations are used after the convolutional layers to reduce the spatial dimensions of the function maps even as retaining the most essential capabilities [1]. once the features had been extracted, the feature maps are flattened right into a one-dimensional vector and then despatched through the fully linked layers. A dropout is used to prevent overfitting through ensuring that the network does no longer memorize the training examples. The softmax activation characteristic is used in the final layer to calculate the opportunity distribution over the five lessons of barcodes [4].

The second one structure is an extension of the CNN version and includes the mixing of an extended short-time period reminiscence (LSTM) network a good way to capture the internal structural dependencies inside the learned characteristic representations. After the convolutional layers are carried out if you want to extract the spatial capabilities from the barcode photograph, the characteristic representation is reshaped in a sequential manner as opposed to being flattened. This reshaped illustration is then surpassed via the LSTM network. LSTM networks are a kind of Recurrent Neural network (RNN) and are in particular effective within the evaluation and detection of internal structural dependencies within the characteristic representation through the usage of memory cells and gate manipulate [2], [3]. The analysis of the internal structural dependencies within the feature representation is based on the relationship between the different parts of the representation. In the proposed model, the Bidirectional LSTM layer is utilized in order to analyze the feature representation in both the forward and backward directions.

After this, the output is passed through dense layers and then the softmax classifier. Both architectures produce the probability value for the classes such as benign barcode, encoded manipulation attack, payload injection attack, structural rotation attack, and adversarial noise attack. From the mentioned classes, the class with the highest probability value is chosen as the output class. It is evident from Fig. 1 that the Pure CNN architecture and CNN-LSTM architecture are compared in order to determine the effectiveness of the inclusion of the concept of sequential learning in the detection of malicious barcode manipulations.

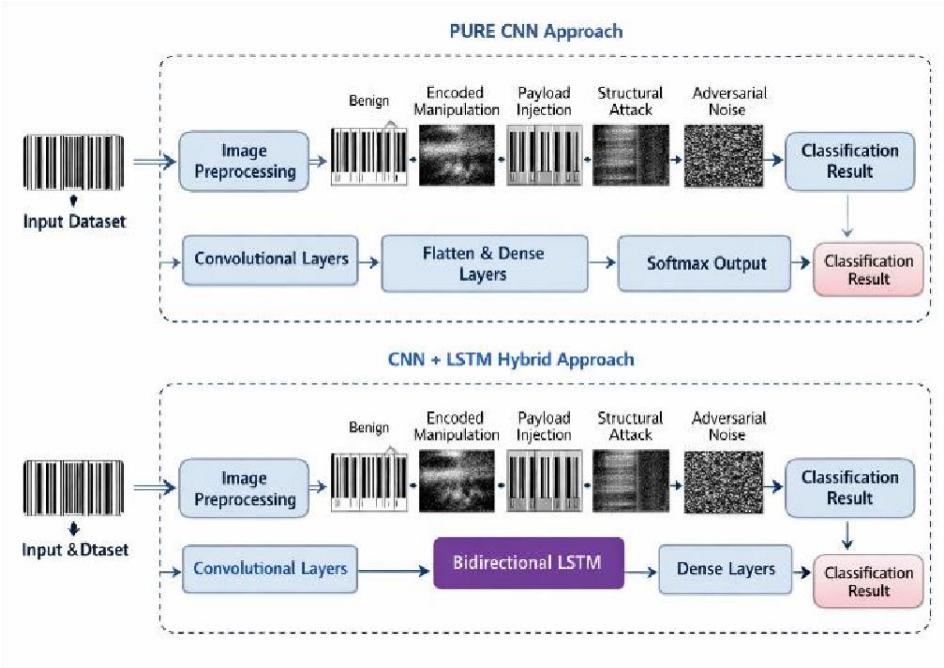

Figure. 1: CNN and CNN-LSTM Flowchart

5. EXPERIMENTAL ANALYSIS

A. Dataset Construction and Data Statistics

The dataset employed in this research was created from a set consisting of 1,461 clean two-dimensional (2D) barcode images. Every clean 2D barcode image was taken as a base image and was transformed into four other types of malicious images through the use of specific attack functions coded in the experimental program. This meant every base image resulted in a total of five images, leading to a multi-class dataset consisting of both clean and tampered images.

The total dataset size was thus calculated as:

$$N_{total} = 16032 \times 5 = 80160 \tag{1}$$

The dataset comprises five classes, each corresponding to a unique condition in the barcode:

- Class 0: Benign barcode condition
- Class 1: Encoded manipulation attack
- Class 2: Payload injection attack
- Class 3: Structural rotation attack
- Class 4: Adversarial noise attack

Prior to education the fashions, the photographs were resized to a defined/fixed spatial decision of 128x128 pixels and are converted to grayscale format. further, pixel values are normalized inside the range [0, 1] to ensure solid education of the neural networks. The normalization method is described as:

$$I_{norm} = \frac{I_{pixel}}{255} \tag{2}$$

represents the normalized pixel value.

Finally, after the preprocessing stage, the dataset is shuffled in random order and split into training and testing sets using an 80:20 split ratio.

The dataset sizes are therefore:

$$N_{train} = 64128,\ N_{test} = 16032 \tag{3}$$

B. Model Configuration and Training Setup

Two architectures of deep learning have been implemented and tested: the Pure Convolutional Neural Network (CNN) and the CNN-LSTM hybrid.

The CNN architecture consists of three blocks of feature extraction layers. In each block, there is a convolution layer, a batch normalization layer, and a max pooling layer. These layers work together to progressively abstract hierarchical spatial features from the barcode image through the learning of discriminative visual patterns within the barcode structure [1], [4].

After the remaining block of the convolution layers, the extracted feature maps are flattened right into a one-dimensional vector, that's then handed through the fully linked layers. further, the dropout layer is used to save you overfitting and improve the generalization overall performance of the CNN. in the final layer, the softmax feature is used as the activation function.

There are a total of 4,288,645 parameters in the CNN model. Out of these parameters, 4,288,197 parameters are trainable parameters, while 448 parameters are non-trainable parameters used in the batch normalization layers.

The second architecture is an extension of the CNN architecture and adds a Bidirectional Long Short-Term Memory (LSTM) network to the model. After convolutional feature extraction, the feature maps are reshaped and fed into a Bidirectional LSTM layer. LSTM networks are capable of extracting long-range dependencies and sequential relationships in structured data, as discussed in [2] and [3].

The total number of parameters in this architecture is 2,356,357, with 2,355,909 trainable parameters and 448 nontrainable parameters.

- Both models were trained using the following configuration:
- Optimizer: Adam
- Batch size: 64
- Number of epochs: 20
- Loss function: categorical cross-entropy

The explicit pass-entropy loss feature used at some stage in training is described as:

$$L = -\sum_{i=1}^{C} y_i \log(\hat{y}_i) \tag{4}$$

The softmax activation function used in the output layer is given by:

$$\hat{y}_i = \frac{e^{z_i}}{\sum_{j=1}^{C} e^{z_j}} \tag{5}$$

C. Training Behavior and Convergence

During the course of the training, both models showed a good rate of convergence as well as a stable pattern of learning. In the case of the CNN model, a high level of accuracy was attained during the early epochs, followed by continuous improvement in the course of training. The level of accuracy in the validation model was attained almost immediately, showing a near-perfect level of classification.

In the case of the hybrid model, a slightly longer period of time was required, owing to the additional processing carried out by the Bidirectional LSTM model.

The total time taken during the course of experimentation, as measured during the training of both models, was as follows:

For the CNN model: ~3040 seconds

For the CNN-LSTM model: ~4057 seconds

The gradual reduction in loss during the course of training indicates a successful convergence of the optimization function.

D. Evaluation Metrics and Numerical Results

After training was completed, both models were evaluated using the test dataset containing 16,032 unseen barcode images. The predicted class label was determined by selecting the class with the maximum softmax probability:

$$\hat{c} = \arg\max_i(\hat{y}_i) \tag{6}$$

The overall classification accuracy was computed using

$$Accuracy = \frac{N_{correct}}{N_{test}} \times 100 \tag{7}$$

The use of this assessment metric, the experimental results potrays to us that the natural CNN model carried out a very last take a look at accuracy of 99.97%, at the same time as the CNN– LSTM hybrid model completed an accuracy of 99.98%.

The classification report displays the precision, recall, and F1-score values are close to 1.00 for each of the five classes, suggesting balanced results for both benign and malicious barcode classes. Also, the confusion matrix shows that the models classify most images correctly with very little misclassification.

These results show that deep learning-based image analysis can identify malicious barcode manipulations with extreme accuracy.

6. DISCUSSION

The experimental results show that both the Pure CNN model and the CNN–LSTM hybrid model are capable of effectively detecting malicious manipulations in 2D barcode images. The results also show that deep learning-based image analysis is capable of successfully detecting structural changes in 2D barcode images without depending on the

content of the barcodes [6], [7]. This is an interesting result since it implies that structural analysis of 2D barcode images is capable of providing effective security analysis for 2D barcode systems.

The training behavior of these models can be observed from the accuracy and loss curves in Fig. 2 and Fig. 3. The accuracy curves in Fig. 2 show that there is rapid increase in accuracy in the initial epochs of training for both of these models. The training accuracy of the Pure CNN model increases sharply in the initial epochs of training and quickly reaches near-perfect classification performance. The CNN–LSTM model also increases in accuracy in subsequent epochs of training, reaching near-perfect performance in classification.

An analysis of Fig. 2 shows that the CNN model's convergence is slightly faster in the initial epochs, which is naturally expected since the model is relatively simpler. On the other hand, the CNN-LSTM model's accuracy curve is smoother in the later epochs of training. The high degree of similarity between the training accuracy and validation accuracy plots for both models is indicative that the models perform well in classifying unseen data, with minimal instances of overfitting [4].

The loss plots in Fig. 3 also show the training process's dynamics for both models. From the loss plots, it is seen that the training loss decreases monotonically with an increase in the epochs for both models, showing that the optimization process is successful in minimizing the classification loss for both models. The CNN-LSTM model's training loss decreases smoothly throughout the training process.

However, it is observed that although there are some fluctuations in the validation loss of the CNN model in the initial epochs of training, these values tend to settle very quickly. These fluctuations are natural in the initial stage of training when the model is trying to adjust its internal parameters to effectively recognize meaningful patterns in images.

The results also show that both models are capable of achieving extremely high classification accuracy for all five classes of barcodes. Class-wise evaluation of the results shows that precision, recall, and F1-score values are very close to 1.00 for all classes. This implies that the models are capable of effectively distinguishing between benign barcodes and different kinds of manipulations performed on them to mimic attacks. There may be some minor variations in recall values in some cases due to structural similarities between benign versions of some classes of barcodes and manipulated versions of some of the attack classes [7], [9].

One of the major factors contributing to the success of the proposed hybrid model is the addition of the Bidirectional Long Short-Term Memory (LSTM) layer. Although the CNN layers are highly effective in the extraction of spatial features in images, the LSTM network is highly efficient in the detection of sequential relationships and dependencies in structured data [2], [3]. In the context of the proposed model dealing with barcode images, the sequential processing of the feature maps allows the model to analyze the internal structural relationships present in the barcode patterns. This allows the model to be more effective in the detection of slight changes introduced in the image by the attacks.

One more advantage of the proposed model is its simplicity in terms of architecture. The dataset is generated based on a set of visual transformations, and the model architectures are not highly complex compared to the complexity present in the architectures of the majority of the deep learning-based models. Despite the moderate complexity in the model architectures, the models achieve extremely high accuracy in the detection task, proving the computational efficiency of the proposed approach

[4].

However, there are some limitations associated with this study. The data used in these experiments are synthetically generated attack data from a base dataset of barcode images. However, in a real-world environment, there are some environmental factors that may influence the barcode images, such as lighting changes, camera perspective, motion blur, and printing quality issues. These environmental issues may bring some extra challenges to the detection of 2D barcode attacks [5].

Overall, it is observed from the experimental results that by integrating CNN for spatial feature extraction and LSTM for sequential learning, it is possible to detect malicious barcodes effectively. From the accuracy and loss curves of the experiment, it is observed that the CNN-LSTM architecture is providing stable convergence with high prediction accuracy for detecting 2D barcode attacks. These results show that deep learning techniques are promising for improving the cybersecurity of 2D barcode systems.

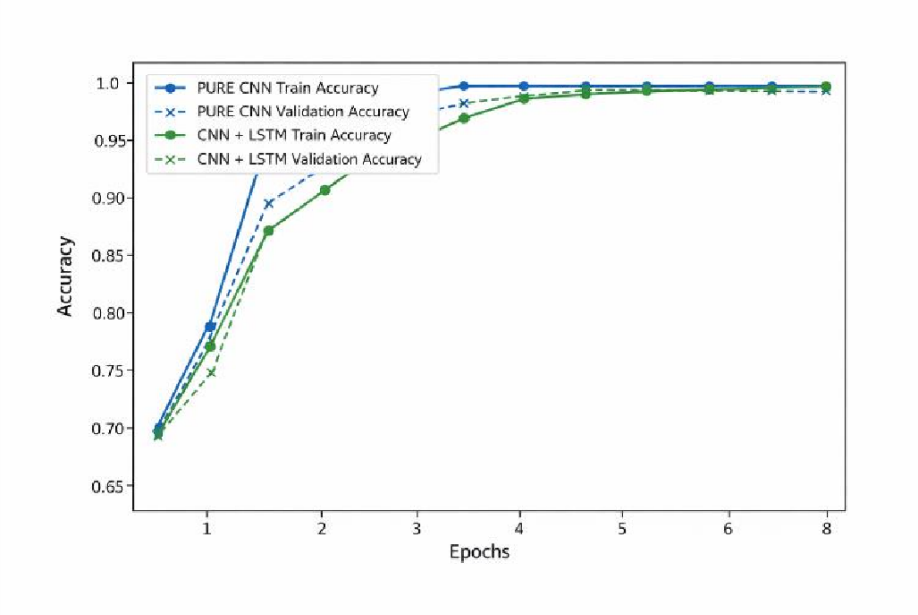

Figure 2: CNN and CNN-LSTM Accuracy Curve

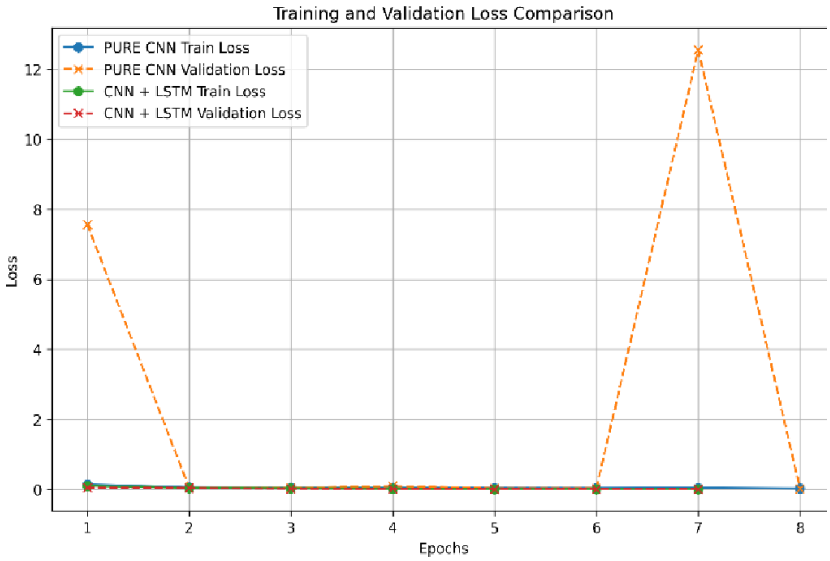

Figure 3: CNN and CNN-LSTM Loss Curve

7. CONCLUSION

This study aimed to solve the problem of detecting malicious manipulations in two-dimensional barcode images from the perspective of visual analysis. Unlike the conventional security approach that relies on the decoded content of the message in the barcode image, the suggested method relies on the visual analysis of the barcode image to

determine whether the image is malicious or not. This is very important in the security domain because cyber attackers can embed malicious links while maintaining the scannability of the image [6], [7].

To solve the problem of detecting malicious manipulations in two-dimensional barcode images from the perspective of visual analysis, two different architectures of deep learning models have been proposed and implemented. These include the Convolutional Neural Network (CNN) and the hybrid CNN– LSTM architectures. in the CNN structure, the focal point is at the extraction of hierarchical spatial functions from the barcode image the use of the convolutional layer of the CNN. In the CNN–LSTM architecture, the focus is on the extension of the CNN architecture by the addition of the Bidirectional Long Short-Term Memory layer [1], [2], [3]. This allows the model to be able to capture the spatial as well as the structural relationships present in the barcode image.

A comprehensive dataset was developed based on the usage of 16,032 clean images of barcodes, wherein each image was transformed into four different versions of manipulated images representing different types of attacks. This resulted in the development of a comprehensive dataset consisting of 80,160 images belonging to five different classes: benign barcode, encoded manipulation attack, payload injection attack, structural rotation attack, and adversarial noise attack. The images have been resized to 128×128 pixels and normalized to the range [0,1] while preserving grayscale. sooner or later, the dataset turned into divided into the schooling and check sets based totally on the eighty:20 ratio. the key characteristics of the dataset are represented in TABLE I.

Table 1

Parameter	**Value**
Total samples after augmentation	80,160
Image size	128 × 128 (grayscale)
Number of classes	5
Training samples	64,128
Testing samples	16,032
Pixel range	0.0 – 1.0

From the experimental results, it is evident that both models were able to achieve extremely high classification performance. The performance results for the standalone CNN version display that the version turned into capable of reap a check accuracy of 99.97%. This proves the effectiveness of the usage of convolutional feature extraction within the classification and identity of structural anomalies within the pictures of barcodes [1], [4].

The precision, recall, and F1-score results were found to be approximately 1.00. This is an indication that the classification performance was well balanced for all the classes in the dataset.

The performance results for the CNN-LSTM model show that the model was able to achieve a slightly higher test accuracy of 99.98%. This proves the effectiveness of the use of the LSTM model in improving the representation capacity of the model.

The LSTM model is able to capture the internal structural relationships in the images of the barcodes. This enables the model to distinguish the differences between the benign and malicious images [2], [3].

The comparative performance results for the two models are shown in Table II.

Table 2

Model	**Accuracy**	**Macro Precision**	**Macro Recall**	**Macro F1score**
CNN	99.97%	1.00	1.00	1.00
CNN– LSTM	99.98%	1.00	1.00	1.00

In addition perception into the class performance can be received from the confusion matrices shown in Fig. 4. Both models achieved perfect classification performance under the binary evaluation setting. In each case, 3,207 benign samples and 12,825 malicious samples were correctly classified, with no misclassified samples. This indicates that the models successfully learned highly discriminative visual features capable of distinguishing between legitimate barcode patterns and maliciously modified ones.

The absence of false positives and false negatives highlights the robustness of the proposed deep learning approach. In practical security systems, minimizing misclassification errors is critical because false alarms can reduce user trust while missed malicious barcodes may expose users to security risks. The results demonstrate that deep learning-based visual inspection can provide a highly reliable mechanism for malicious barcode detection.

By combining spatial feature extraction through convolutional layers with structural sequence learning through LSTM networks, the proposed framework achieves accurate and stable classification performance. Overall, the results confirm that deep learning-based visual analysis offers a promising solution for strengthening the security of real-world barcode systems, particularly in scenarios where analyzing decoded barcode content alone is insufficient for detecting malicious intent [6], [7], [10], [11].

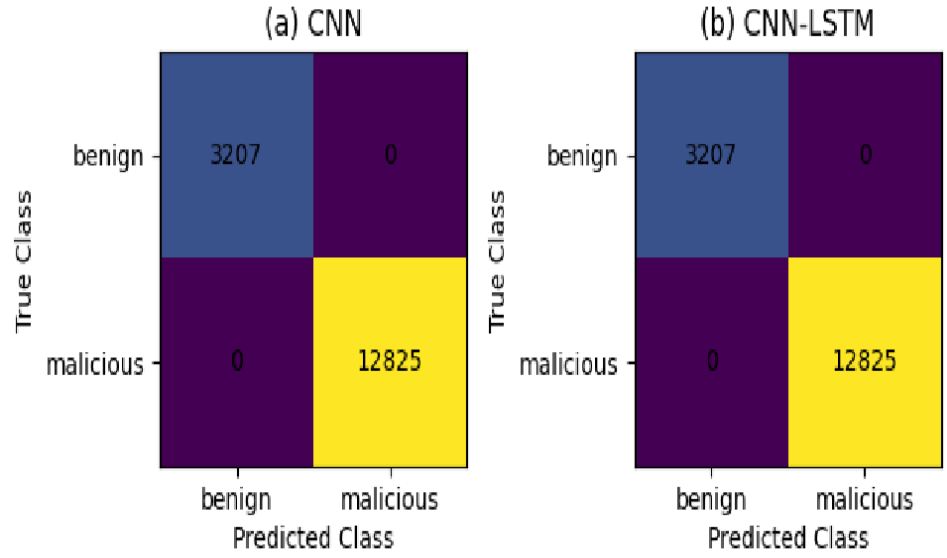

Figure 4: Confusion Matrix

8. ACKNOWLEDGMENT

The author Pranav Pradeep Menon wishes to express his sincere appreciation to the professors of VIT Bhopal University, who were very significant for the progress of the work. He is very appreciative to his supervisor, Devraj Vishnu, for the constant support he provided with patience. Special appreciation goes to him for the technical advice he offered throughout the planning, implementation, and documentation of the study. He is of immense significance with regard to the orientation of the research.

Another major aspect, which the author shall never forget to be deeply revered, is to express the referees, Narottam Das Patel and Praveen Lalwani, a heart-felt thank you for the comments, criticism, and suggestions that are of immense worth to clarify the paper, add technical depth to the work, and make the whole reading a breeze, etc.

Finally, Pranav Pradeep Menon would like to take this opportunity to thank all the teachers and staff of VIT Bhopal University who, directly or indirectly, have contributed incessantly to the successful completion of this study through constant inspiration, academic support, and providing good conditions for learning.

REFERENCES

1. K. He, X. Zhang, S. Ren, and J. Sun, “Deep residual learning for image recognition,” in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016, pp. 770–778. https://ieeexplore.ieee.org/document/7780459
2. R. C. Staudemeyer, “Long Short-Term Memory recurrent neural network,” *Procedia Computer Science*, vol. 125, pp. 332–339, 2018. https://www.sciencedirect.com/science/article/pii/S1877050917328557? via%3Dihub
3. S. Hochreiter, R. C. Staudemeyer, and J. Schmidhuber, “Understanding LSTM: A tutorial into Long Short-Term Memory recurrent neural networks,” arXiv preprint, 2019. https://arxiv.org/pdf/1909.09586
4. Y. LeCun, Y. Bengio, and G. Hinton, “Deep learning,” *Nature*, vol. 521, no. 7553, pp. 436–444, 2015. https://www.nature.com/articles/nature14539
5. M. Abdellatif, M. Abdel-Basset, M. Soliman, and H. Tolba, “A survey of barcode and QR code security vulnerabilities and their countermeasures,” *Information Sciences*, vol. 632, pp. 17–46, 2024. https://www.mdpi.com/2624-800X/2/4/41
6. R. Wudhikarn, P. Charoenkwan, and K. Malang, “Deep learning in barcode recognition: A systematic literature review,” *IEEE Access*, vol. 10, ppl. 1-18, 2022. https://ieeexplore.ieee.org/abstract/document/9681298
7. M. Toğaçar and B. Ergen, “Processing 2D barcode data with metaheuristic based CNN models and detection of malicious PDF files,” *Applied Soft Computing*, vol. 154, Art. no. 111358, 2024. https://www.sciencedirect.com/science/article/abs/pii/S15684946240049 64
8. M. Chemmakha, O. Habibi, and M. Lazaar, “Improving machine learning models for malware detection using embedded feature selection method,” *Journal of Information Security and Applications*, vol. 74, Art. no. 103524, 2023. https://www.sciencedirect.com/science/article/pii/S2405896322008072? via%3Dihub
9. F. Rustam, I. Ashraf, A. D. Jurcut, A. K. Bashir, and Y. B. Zikria, “Malware detection using image representation of malware data and transfer learning,” *Journal of Information Security and Applications*, vol. 66, Art. no. 103166, 2022. https://www.sciencedirect.com/science/article/abs/pii/S07437315220021 18?via%3Dihub
10. S.-Y. Yang, H.-C. Jan, C.-Y. Chen, and M.-S. Wang, “CNN-based QR code reading of package for unmanned aerial vehicle,” *Sensors*, vol. 23, no. 10, Art. no. 4707, 2023. https://www.mdpi.com/14248220/23/10/4707
11. B. B. Traore, B. Kamsu-Foguem, and F. Tangara, “Deep convolution [12] AI Generated Barcode Images/Masks Dataset neural network for image recognition,” *Ecological Informatics*, vol. 48, https://www.kaggle.com/datasets/bd2927/ai-generated-barcode- pp. 257–268, 2018. imagesmasks-dataset