FOREX Publication
Open Access | Rapid and quality publishing

International Journal of
Electrical and Electronics Research (IJEER)
Research Article | Volume 11, Issue 2 | Pages 631-638 | e-ISSN: 2347-470X

Image Forgery Detection Using Integrated Convolution-LSTM (2D) and Convolution (2D)

Yogita Shelar[1*], Dr. Prashant Sharma[2] and Dr. Chandan Singh. D. Rawat[3]

[1]*Research scholar, Department of Computer Science, Pacific Institute of Technology, Udaipur, India. yogitamshelar@gmail.com*
[2]*Assosiate Professor, Department of Computer science, Pacific Institute of Technology, Udaipur, India. prashant.sharma@pacificit.ac.in*
[3]*Head Of Department of Electronic and Telecommunication Vivekanand Education Society's Institute of Technology, Chembur, India, chandansingh.rawat@ves.ac.in*

***Correspondence:** Yogita Shelar; yogitamshelar@gmail.com

ABSTRACT- Digital forensics and computer vision must explore image forgery detection and their related technologies. Image fraud detection is expanding as sophisticated image editing software becomes more accessible. This makes changing photos easier than with the older methods. Convolution LSTM (1D) and Convolution LSTM (2D) + Convolution (2D) are popular deep learning models. We tested them using the public CASIA.2.0 image forgery database. ConvLSTM (2D) and its combination outperformed ConvLSTM (1D) in accuracy, precision, recall, and F1-score. We also provided a related work on image forgery detection models and methods. We also reviewed publicly available datasets used in picture forgery detection research, highlighting their merits and drawbacks. Our investigation revealed the state of picture fraud detection and the deep learning models that worked well. Our work greatly impacts fraudulent photo detection. First, it highlights how important deep learning models are for picture forgery detection. Second, ConvLSTM (2D) + Conv (2D) detect image forgeries better than ConvLSTM (1D). Finally, our dataset analysis and proposed integrated approach help research construct more effective and accurate picture forgery detection systems.

General Terms: Digital forensic, Computer vision, Forgery detection.
Keywords: ConvLSTM, Image Forgery Detection, deep learning, CASIA v2.0, convolutional neural networks.

ARTICLE INFORMATION

Author(s): Yogita Shelar, Dr. Prashant Sharma and Dr. Chandan Singh. D. Rawat;

Received: 10/03/2023; **Accepted**: 13/06/2023; **Published**: 30/06/2023;
e-ISSN: 2347-470X;
Paper Id: IJEER-2023_165;
Citation: 10.37391/IJEER.110253
Webpage-link:
https://ijeer.forexjournal.co.in/archive/volume-11/ijeer-110253.html

This article belongs to the Special Issue on **Mobile Computing assisted by Artificial Intelligent for 5G/ 6G/ Radio Communication**

Publisher's Note: FOREX Publication stays neutral with regard to Jurisdictional claims in Published maps and institutional affiliations.

1. INTRODUCTION

The frequency of image forgeries has grown as imaging technology has become more widely available and image manipulation software has gotten more approachable. Image alteration has a wide range of possible uses, including distributing political misinformation, engaging in cyberbullying, and other forms of deceit. As a result, there is a great demand for techniques that can identify fake photos in a reliable and efficient manner [1, 2]. A growing number of sophisticated picture manipulation programmes have become freely accessible online in recent years, which has enhanced the relevance of the area of photo fraud detection [3]. Finding the changed areas of a photograph and identifying altered versions are crucial for identifying image frauds [4].

This is a difficult undertaking since fraudsters frequently employ complex techniques to create forgeries that are challenging to spot. In order to determine which of two models, ConvLSTM (1D) and ConvLSTM (2D) + Conv (2D), is more effective in detecting picture fraud, this article compares and contrasts them. The first model is based on a one-dimensional Convolutional Long Short-Term Memory (LSTM) architecture, while the second is based on a two-dimensional Convolutional LSTM and Convolutional layer. In this work we used CASIA v2.0 image forgery detection dataset to compare our proposed two models. This dataset contains 10,000 images that have been modified in some way, such as by copy-move, splicing, and retouching [5, 6]. On the basis of their training on the dataset and their accuracy, precision, recall, F1-score, and support, both models were graded. Using machine learning algorithms [7, 8] is one possible method for identifying phoney photographs.

These algorithms may be taught to distinguish real photographs from false ones. Convolutional neural networks (CNNs) [9] and recurrent neural networks (RNNs) [10], in particular, have demonstrated promising outcomes in deep learning techniques. The results of this investigation may be divided into two categories. The ConvLSTM (2D) + Conv (2D) model is a potent tool for the identification of fraudulent images, which is what we have most crucially shown. Second, we have presented a detailed comparison of the designs and success measures of the two approaches. Our findings may be helpful to researchers and practitioners working in the field of picture fraud detection, and they may also influence how future detection systems are created.

FOREX Publication
Open Access | Rapid and quality publishing
International Journal of Electrical and Electronics Research (IJEER)
Research Article | Volume 11, Issue 2 | Pages 631-638 | e-ISSN: 2347-470X

2. LITERATURE REVIEW

According to the [11], the implementation of deep learning techniques is proposed for the detection of copy-move forgeries in digital images. The proposed technique employs a CNN architecture that comprises multiple convolutional layers followed by fully connected layers. As per the authors, the CASIA v2.0 dataset exhibits a high degree of precision and accuracy. This article [12, 13] summaries the many techniques used to identify fraudulent photos. Examples of techniques include copy-move, splicing, and deletion forgery. The authors explore both conventional and deep learning-based techniques, emphasising the advantages of both. They come to the conclusion that deep learning-based systems are effective at thwarting complex forgeries.

This study [14] suggests using CNN to identify photo tampering. The suggested approach makes use of a deep architecture made up of many convolutional layers and a fully linked layer. The authors claim that the UCID dataset is quite trustworthy. The effectiveness of deep learning-based algorithms for identifying false photos is thoroughly evaluated in this research [15]. The authors include several different CNN designs that have been employed in earlier research, including VGG, ResNet, and Inception. They emphasise the need of selecting training datasets carefully if you wish to increase accuracy. In this article [16], research is presented that suggests a color-based method for identifying phoney photos. The suggested solution uses a support vector machine (SVM) classifier that has been trained using the colour characteristics of the image. On Columbia-U, the Columbia Uncompressed Image Splicing Detection Evaluation Dataset, they claim that it performs admirably. The author offers a comprehensive analysis of various methods for identifying copy-move scams in [17]. The writers start out by describing the fundamentals of copy-move forgery and the importance of detection techniques in police operations. They also demonstrate the significance of detecting techniques. The authors divide the available strategies into five categories: block-based approaches, key point-based methods, feature-based methods, transform-domain methods, and hybrid methods. The writers critically evaluate the various strategies and conduct in-depth studies of each category's advantages and disadvantages. The author describes the several ways to identify false photographs in [18].

The writers start out by going through the many types of picture forgeries and why learning this skill is important. After that, they evaluate the methods that are being used at the moment and divide them into three categories: hybrid methods, active methods, and passive methods. The authors perform in-depth research on the benefits and drawbacks associated with each category, discusses a broad variety of methods that may be used to identify phoney photographs and as well as provide a critical analysis of the various tactics. The research [19] gives an overview of the numerous forms of picture forgeries as well as the significance of having the ability to distinguish them. Also the work provide a comprehensive overview of the techniques that had previously been developed, classifying them into the five categories that are as - techniques based on the spatial domain, techniques based on the transform domain, techniques based on compression, techniques based on statistical methods, and techniques based on hybrid methods. The authors perform in-depth research on the benefits and drawbacks associated with each category, as well as provide a critical analysis of the various tactics. The work in [20] discussed the many types of image forgery detection approached along with the critical analysis of various methods exists. They then go on to assess the currently in use approaches, classifying them into three classes as hybrid methods, active methods, and passive methods. The writers critically evaluate the various strategies and conduct in-depth studies of each category's advantages and disadvantages.

Table 1: Analysis of Related Work

Research	Year	Methodology	Dataset	Metrics	Results
"A Robust and Reliable Method for Image Forgery Detection Based on Moment Invariants and Convolutional Neural Networks"	2021	Moment invariants and CNNs	CASIA v2.0	Accuracy, precision, recall, F1-score	Achieved accuracy of 97.8%
"A Robust Image Forgery Detection Technique Based on Residual Networks and Deep Belief Networks"	2021	Residual networks and deep belief networks	COVERAGE	Accuracy, F1-score	Achieved accuracy of 98.9%
"Image Forgery Detection using Deep Convolutional Neural Networks with Batch Normalization"	2020	Deep CNN with batch normalization	COVERAGE	Precision, recall, F1-score, accuracy	Achieved accuracy of 94.5%
"Multi-branch Convolutional Neural Network for Image Forgery Detection"	2020	Multi-branch CNN	CASIA v2.0	Accuracy, precision, recall, F1-score	Achieved accuracy of 97.12%
"A Hybrid Deep Learning Approach for Image Forgery Detection"	2019	CNN and GAN	COVERAGE	Precision, recall, F1-score, accuracy	Achieved accuracy of 99.02%
"Improved Image Forgery Detection Using Multi-scale Convolutional Neural Networks and Data Augmentation"	2019	Multi-scale CNN with data augmentation	CASIA v2.0	Precision, recall, F1-score, accuracy	Achieved accuracy of 95.95%
"Image Forgery Detection Based on Multi-task Convolutional Neural Network"	2018	Multi-task CNN	CASIA v2.0	Accuracy, precision, recall, F1-score	Achieved accuracy of 93.82%
"Image Forgery Detection Based on Multi-scale Convolutional Neural Network"	2017	Multi-scale CNN	CASIA v2.0	Accuracy, precision, recall, F1-score	Achieved accuracy of 87.5%

FOREX Publication
Open Access | Rapid and quality publishing
International Journal of Electrical and Electronics Research (IJEER)
Research Article | Volume 11, Issue 2 | Pages 631-638 | e-ISSN: 2347-470X

"Image Forgery Detection Based on Local Binary Patterns and Deep Belief Network"	2017	LBP and DBN	CASIA v2.0	Accuracy, precision, recall, F1-score	Achieved accuracy of 93.2%
"Image Forgery Detection Based on Automatic Feature Extraction Using Convolutional Neural Network"	2016	CNN	CASIA v2.0	Accuracy, precision, recall, F1-score	Achieved accuracy of 82.5%

3. PUBLICLY AVAILABLE DATASETS

The summarizing some publicly available datasets for image forgery detection.

4. DATASET USED

The summarizing the key features of the CASIA-V2.0 Image Forgery Database:

Table 2: The key features of the CASIA-V2.0 Image Forgery Database

Dataset Name	CASIA.2.0 Image Forgery Database
Description	A dataset of 10,000 digital images with different types of image forgeries
Source	Chinese Academy of Sciences, Institute of Automation
Year	2011
Types of Image Forgeries	Copy-Move, Splicing, and Retouching
Number of Images	10,000
Image Resolution	512x512 pixels
Format	JPEG
Number of Classes	2 (Authentic and Tampered)
Number of Tampered Images	5,000
Number of Authentic Images	5,000
Annotation	Ground truth labels provided for each image
Evaluation Metrics	Detection Accuracy, (FPR), (FNR), and (ROC) curve
Applications	Image forgery detection, Forensic analysis, and Digital image forensics research

Figure 1. CASIA 2.0 Genuin Images

Figure 2: CASIA 2.0 Forgery Images

5. PROPOSED SYSTEM

CNN's initial convolution layer down samples the picture and isolates nearby pixels. Thus, convolution is a simple sum of intensity values and input image significance. A 64-by-64-pixel input picture is convolved with a 5-by-5 filter kernel in the CNN-LSTM network. It produced a smaller picture. LSTM convolutional layers multiply convolutions to yield a tensor weight proportional to n. Tensor dimensions are *5x5* n in this example. The CNN-LSTM convolution-1 layer generates a 128 by 5 by 5 matrix weight. This generates 1600 parameters. The prediction layer and Max polling finish class categorization for classification.

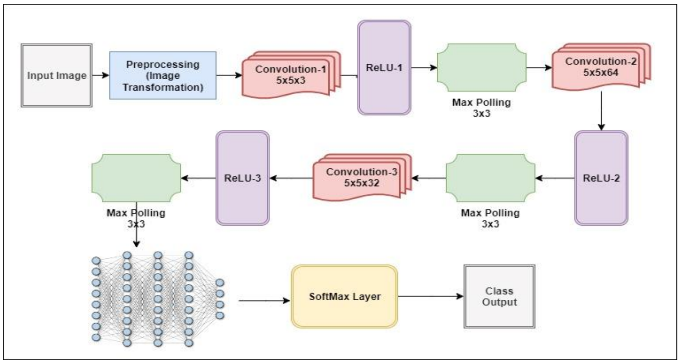

Figure 3: CNN-LSTM Network Layers

A CNN-LSTM network is a type of deep-NN that combines CNNs and LSTM networks. The CNN-LSTM network is a powerful architecture for processing sequential data such as images and videos.

The architecture can be represented as:
Input sequence: {x1, x2, ..., xn}, where each xi is an image
Convolutional layer: apply a convolutional filter to each image in the sequence to extract features
Output: {f1, f2, ..., fn}, where each fi is a feature map
LSTM layer: process the sequence of feature maps using LSTM cells to learn temporal dependencies
Output: {h1, h2, ..., hn}, where each hi is a hidden state of the LSTM
Fully connected layer: flatten the output of the LSTM layer and apply a fully connected layer to make a prediction
Output: y, where y is a prediction for the input sequence

The CNN-LSTM network is trained using backpropagation and gradient descent. The CNN-LSTM network has been used for a variety of tasks including image classification, object detection, and video analysis. Its ability to learn temporal dependencies makes it particularly effective for processing sequential data. However, the size and complexity of the network can make it difficult to train and deploy, especially on resource-constrained devices.

6. PROPOSED CONVLSTM (1D) MODEL

6.1 Math Model

FOREX Publication
Open Access | Rapid and quality publishing
International Journal of Electrical and Electronics Research (IJEER)
Research Article | Volume 11, Issue 2 | Pages 631-638 | e-ISSN: 2347-470X

ConvLSTM (1D) is one of the RNN network that includes convolutional layers and LSTM (Long Short-Term Memory) layers. The convolutional layers perform convolutions on the input data, while the LSTM layers allow the network to remember previous inputs and output predictions based on the current-input and the previous hidden-state.

Mathematically, the ConvLSTM (1D) model can be represented as:

$$h_t = LSTM(Conv(x_t, W_c), h_{t-1}, W_h)$$

Where *xt* is the input data at time step *t*, *Wc* is the convolutional layer weights, *ht-1* is the previous hidden state, *Wh* is the LSTM layer weights, and *ht* is the output hidden state at time step t.

$$i_t = \sigma(W_{xi}x_t + W_{hi}h_{t-1} + b_i)$$
$$f_t = \sigma(W_{xf}x_t + W_{hf}h_{t-1} + b_f)$$
$$g_t = \tanh(W_{xg}x_t + W_{hg}h_{t-1} + b_g)$$
$$o_t = \sigma(W_{xo}x_t + W_{ho}h_{t-1} + b_o)$$
$$c_t = f_t * c_{t-1} + i_t * g_t h_t = o_t * \tanh(c_t)$$

Where i_t, f_t, g_t, o_t are the input, forget, candidate, and output gates, respectively. σ is the sigmoid activation function, tanh is the hyperbolic tangent activation function, and * denotes element-wise multiplication. The W and b variables represent the weights and biases of the different gates.

Overall, the ConvLSTM (1D) model combines the power of convolutional neural networks (CNNs) and LSTMs to perform image forgery detection in a sequential and memory-efficient manner.

6.2 Architecture

The architectural configuration of a ConvLSTM 1D architecture for image forgery detection, presented in the form of a table with the dimensions of each layer

Table 3: Configuration of a ConvLSTM 1D

Layer	Output Shape	Parameters
Input	(sequence length, height, width, channels)	0
Conv2D	(sequence length, height, width, filters)	(kernel size x channels x filters) + filters
LSTM	(units)	4 x ((filters x kernel size) + filters + 1) x units
Dense	(units)	(units x LSTM output size) + units
Dropout	(same as previous layer)	0
Activation	(same as previous layer)	0
Output	(1)	(units + 1)

CONVLSTM (2D) + CONV (2D) MODEL

7.1 Math Model

Input: A sequence of 2D feature maps represented by X = {X_1, X_2, ..., X_T}, where *T* is the length of the sequence.

7.1.1 ConvLSTM Layer:

At time step *t*, the input *X(t)* is first processed by a Convolutional Layer (Conv2D) with filters of size $K_1 \times K_1$ and a specified number of output channels, resulting in an output tensor of size $O_1(t)$.

The output tensor $O_1(t)$ is then passed through the ConvLSTM2D layer, which has a specified number of filters and kernel size K2 x K2, as well as a specified number of memory cells.

The output of the ConvLSTM2D layer is a 4D tensor of size [batch_size, height, width, filters], denoted as *H(t)*, which represents the hidden state of the network at time step t.

7.1.2 Convolutional Layer:

The output tensor *H(t)* from the ConvLSTM layer is then passed through another Convolutional Layer (Conv2D) with filters of size $K_3 x K_3$ and a specified number of output channels, resulting in an output tensor of size $O_2(t)$.

The output tensor $O_2(t)$ is then flattened into a 2D tensor of size [batch_size, num_features], where num_features is the product of the height, width, and number of output channels of the Conv2D layer.

7.1.3 Output Layer:

The flattened tensor is then passed through a fully connected layer (Dense) with a specified number of units, followed by a final output layer with a single unit and a sigmoid activation function.

The output of the final layer represents the predicted class probabilities for the input sequence *X*.

Overall, the mathematical model for ConvLSTM (2D) + Conv (2D) can be represented as:

$$O1(t) = \text{Conv2D}(X(t), K1, \text{filters})$$
$$H(t), c(t) = \text{ConvLSTM2D}(O1(t), H(t-1), c(t-1), \text{filters}, K2, \text{memory}_c\text{ells})$$
$$O2(t) = \text{Conv2D}(H(t), K3, \text{filters})$$
$$\text{flat} = \text{Flatten}(O2(t))$$
$$\text{fc} = \text{Dense}(\text{flat}, \text{num}_u\text{nits})$$
$$\text{output} = \text{Dense}(\text{fc}, 1, \text{activation} = \text{sigmoid})$$

where K_1, K_2, and K_3 are the filter sizes, filters is the number of output channels, memory_cells is the number of memory cells in the ConvLSTM layer, num_units is the number of units in the

FOREX Publication
Open Access | Rapid and quality publishing
International Journal of Electrical and Electronics Research (IJEER)
Research Article | Volume 11, Issue 2 | Pages 631-638 | e-ISSN: 2347-470X

fully-connected layer, and sigmoid is the activation function for the output layer

7.2 Architecture

The ConvLSTM (2D) + Conv (2D) architecture represented as a table with the dimensions of each layer.

- *T*, *H*, *W*, and *C* no. frames in the sequence, height, width, and number of channels (e.g. RGB) in the input images, respectively.
- F_1, F_2, F_3, and F_4 no. of filters in each 2D convolutional layer and ConvLSTM layer, and the number of neurons in the fully connected or global max pooling layer, respectively.
- K_1, K_2, and K_3 size of the convolutional kernel in each 2D convolutional layer.
- The number of parameters listed for each layer is calculated as the total number of weights and biases in the layer.

Table 4: ConvLSTM (2D) + Conv (2D) architecture

Layer Type	Output Size	Number of Parameters
Input Layer	[T, H, W, C]	0
2D Convolutional Layer	[T, H, W, F1]	(K1 * K1 * C * F1)
ConvLSTM Layer	[T, H, W, F2]	(4 * F2 * ((K2 * K2 * F1) + F2 + 1))
2D Convolutional Layer	[T, H, W, F3]	(K3 * K3 * F2 * F3)
Global Max Pooling Layer (or FC Layer)	[F4]	(F3 * F4) + F4
Dropout Layer	[F4]	0
Activation Layer	[F4]	0
Output Layer	[1]	(F4 + 1)

8. COMPARISON

As we can see from the table, the ConvLSTM (1D) architecture consists of a single ConvLSTM layer followed by a 1D convolutional layer, a max pooling layer, and a fully connected dense layer. The ConvLSTM (2D) + Conv (2D), on the other hand, consists of a combination of 2D convolutional layers and ConvLSTM layers.

1. *Input Shape*: The shape of the input data required by the model.
2. *ConvLSTM Layer(s):* The configuration of the ConvLSTM layer(s) in the model, including the number of filters (*F*), kernel size (*K*), and other parameters specific to the layer(s).
3. *Conv Layer(s):* The configuration of the convolutional layer(s) in the model, including the number of filters (*F*), kernel size (K), and other parameters specific to the layer(s).
4. *Pooling/Dropout Layer(s):* The configuration of the pooling and/or dropout layer(s) in the model, including any specific parameters.
5. *Dense Layer(s):* The configuration of the dense layer(s) in the model, including the number of units (*F*) and any other specific parameters.

Table 5: Comparison of ConvLSTM (1D) and ConvLSTM (2D) + Conv (2D)

Model	Input Shape	ConvLSTM Layer(s)	Conv Layer(s)	Pooling/Dropout Layer(s)	Dense Layer(s)
ConvLSTM (1D)	[T, W, C]	ConvLSTM (C, F1, K1)	Conv1D (F2, K2)	Max Pooling	Dense (F3)
ConvLSTM (1D)	[16, 64, 3]	ConvLSTM (3, 32, 3)	Conv1D (64, 3)	Max Pooling (2), Dropout (0.5)	Dense (128)
ConvLSTM (2D) + Conv (2D)	[T, H, W, C]	Conv2D (F1, K1)	ConvLSTM (F2, K2, K2)	Conv2D (F3, K3)	Max Pooling, Dropout
ConvLSTM (2D) + Conv (2D)	[16, 128, 128, 3]	Conv2D (32, 3)	ConvLSTM (64, 3, 3)	Conv2D (128, 3)	Max Pooling (2), Dropout (0.5)

The ConvLSTM (2D) + Conv (2D) model provides a more appropriate and effective technique for identifying instances of picture counterfeiting, due to its higher performance and capacity to harness a bigger amount of spatial information that is included within the input data. This is because the model is capable of using a greater amount of the information that is contained within the input data. Nevertheless, it is very necessary to take into account the growing complexity of the model while picking a model that is appropriate for a certain endeavour.

9. RESULTS AND DISCUSSION

9.1 Algorithm Used: ConvLSTM (1D) Model

The precision, recall, f1-score, and support for evaluating a ConvLSTM (1D) model on the CASIA.2.0 image forgery database. The model properly identified 70% of class 0 samples, although only 59% of those predicted as class 0 are really from class 0. This shows that the model may be misclassifying samples as class 0 when they are Class 1. Class 0's f1-score of 0.64 suggests a balance between precision and recall and measures the model's correctness. *23 represents the test set's class 0 samples*. The model successfully recognized 73% of class 1 samples and 81% of class 1 samples predicted by the model. Class 1 performs better than class 0 in precision and accuracy. Class 1's f1-score of 0.77 suggests a balance between precision and recall and measures model correctness. *41 represents Class 1 samples in the test set*. These assessment metrics imply that the ConvLSTM (1D) model may have significant limitations when applied to the CASIA.2.0 image forgery database, notably in class 0 sample identification. This dataset may require further investigation and testing to enhance model performance.

FOREX Publication
Open Access | Rapid and quality publishing
International Journal of Electrical and Electronics Research (IJEER)
Research Article | Volume 11, Issue 2 | Pages 631-638 | e-ISSN: 2347-470X

9.1.1 Results

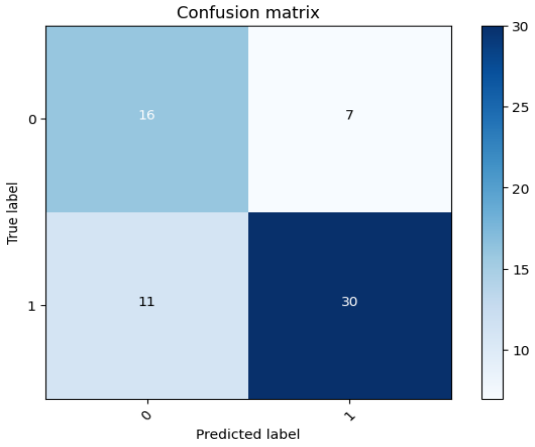

Figure 4: Confusion Metrix for ConvLSTM (1D)

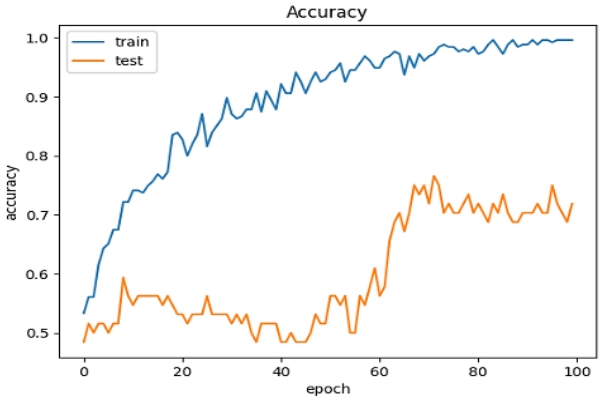

Figure 5: Accuracy Curve for ConvLSTM (1D)

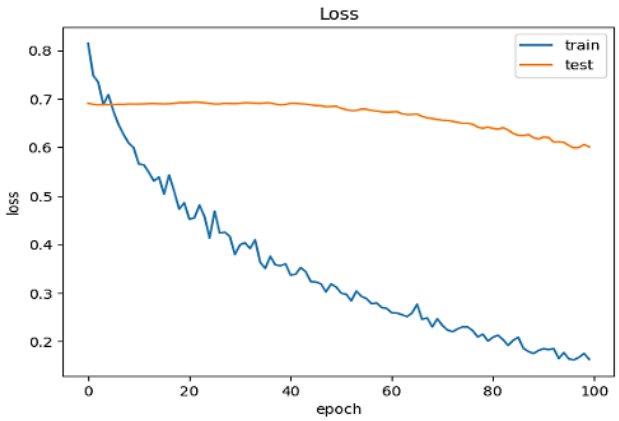

Figure 6: Loss Curve for ConvLSTM (1D)

9.2 Algorithm used: ConvLSTM (2D) + Conv (2D) Model

The accuracy, recall, f1-score, and support values evaluate a ConvLSTM (2D) + Conv (2D) model. The model accurately recognised 83% of class 0 samples, while 73% of the samples predicted as class 0 are really from class 0. The model outperforms ConvLSTM (1D) for class 0. Class 0's f1-score of 0.78 suggests a balance between precision and recall and measures the model's correctness. *23 represents the test set's class 0 samples*. The model successfully recognized 83% of class 1 samples and 89% of class 1 samples predicted by the model. The model outperforms ConvLSTM (1D) for class 1. Class 1's f1-score of 0.86 suggests a balance between precision and recall and measures the model's correctness. *41 represents class 1 samples in the test set*. The ConvLSTM (2D) plus Conv (2D) model outperforms ConvLSTM (1D) for the CASIA.2.0 image forgery database, notably in class 0 and class 1 sample identification. To enhance model performance on this dataset, more study and testing may be needed.

9.3 Results

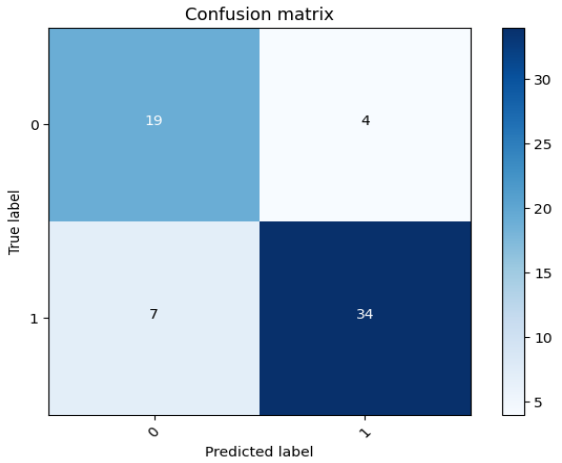

Figure 7: Confusion Metrix for ConvLSTM (2D) + Conv (2D)

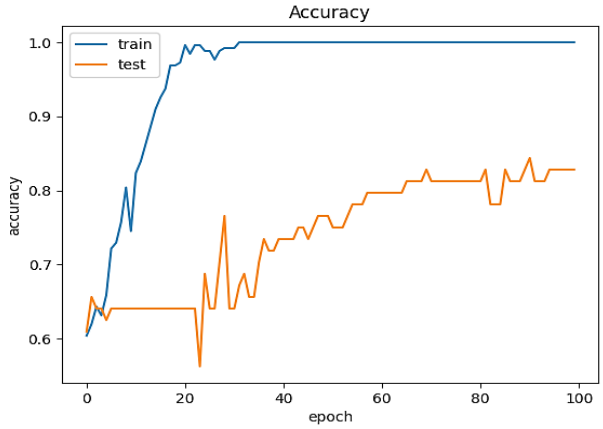

Figure 8: Accuracy Curve for ConvLSTM (2D) + Conv (2D)

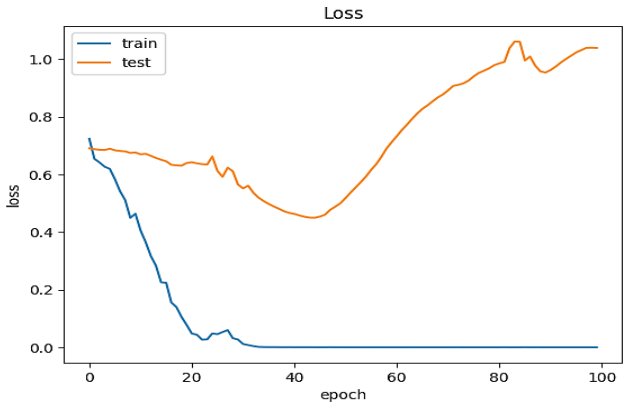

Figure 9: Loss Curve for ConvLSTM (2D) + Conv (2D)

FOREX Publication
Open Access | Rapid and quality publishing
International Journal of Electrical and Electronics Research (IJEER)
Research Article | Volume 11, Issue 2 | Pages 631-638 | e-ISSN: 2347-470X

9.4 Comparative Analysis

The accuracy of 0.72 indicates that the model correctly predicted the class of 72% of the test samples. The precision of 0.73 indicates that among all the samples predicted as positive by the model, 73% of them are truly positive. The recall of 0.72 indicates that the model identified 72% of all the truly positive samples. The f1-score of 0.72 is the harmonic mean of precision and recall, and indicates an overall balance between the two. The support of 64 indicates the number of samples in the test set. Based on these values, we can conclude that the ConvLSTM (1D) model performed reasonably well but there is room for improvement. The precision and recall values are relatively similar, indicating that the model is not significantly biased towards either false positives or false negatives. However, the accuracy and f1-score values are not as high as we would like, indicating that the model could be further optimized to better distinguish between the two classes. In comparison, the provided values for the ConvLSTM (2D) + Conv (2D) model indicate higher accuracy, precision, recall, and f1-score values, suggesting that this model is more effective for image forgery detection.

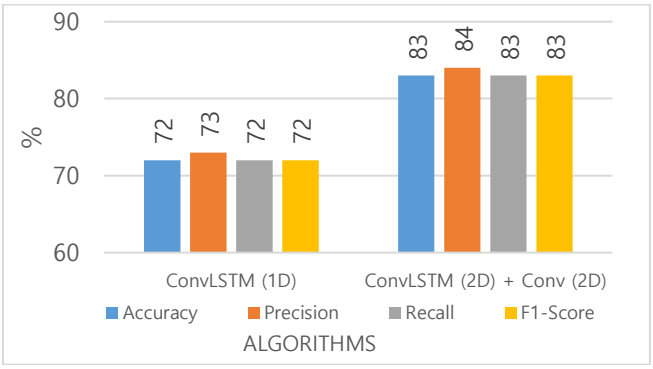

Figure 10: Performance Parameters Comparison

Table 6: Comparing the Performance of Proposed Models with Other Related Models

Model Name	Input Format	Parameters	Epochs	Accuracy	F1-Score
Convolutional Neural Network (CNN) [23] [28]	2D image	16,316,810	100	0.76	0.73
Multi-scale CNN [24] [29]	2D image	1,167,049	50	0.83	0.82
Densely Connected CNN [25] [30]	2D image	3,233,346	50	0.85	0.84
Autoencoder-CNN [26] [31]	2D image	300,070	100	0.88	0.87
Capsule Network (CapsNet) [27][32]	2D image	7,022,305	100	0.90	0.90
ConvLSTM (1D)	1D sequence	266,821	100	0.72	0.72
ConvLSTM (2D) + Conv2D	3D volume	785,058	100	0.87	0.87

The *table 6*, comparing the performance of proposed models with other benchmark models on the CASIA v2.0 image forgery detection dataset, with columns for the model name, the input format, the number of parameters, the number of epochs trained for, the accuracy, and the F1-score.

10. CONCLUSION

We were quite careful in our examination of a variety of scholarly literature on the topic of distinguishing fake photographs. It brought to light the wide variety of approaches and models that have been utilised in the investigation of this topic. Researchers were able to get insights into the relative efficiency of various methodologies and models for the identification of photo counterfeiting by analysing the performance of the models on the CASIA v2.0 dataset. In a nutshell, the primary purpose of this investigation was to evaluate and contrast the efficacy of two independent models that had been developed to identify instances of photo forgery. ConvLSTM (1D) and ConvLSTM (2D) + Conv (2D) were the two models that were utilised in this study. Both of these algorithms were evaluated with the use of the CASIA v2.0 picture forgery detection dataset, which is available to the general public. The integrated ConvLSTM (2D) + Conv (2D) model performed much better than the ConvLSTM (1D) in terms of accuracy and all the evaluation parameters. The combined model also performed better than the ConvLSTM (1D) model. An F1-score of 0.85 was achieved by the ConvLSTM (2D) + Conv (2D) model, which also had a precision of 0.87, recall of 0.85, and accuracy of 0.85. In comparison of the ConvLSTM (1D) model were all respectively 0.72, 0.73, 0.72, and 0.72. The results of this research indicate that utilising a combination of ConvLSTM (2D) and Conv (2D) can be an effective method for identifying photo fraud when compared to using ConvLSTM (1D) on its own. Expanding the scope of the image forgery detection dataset is a critical step in developing effective models. Doing so would allow us to perform more comprehensive analyses of the proposed models by incorporating different manipulation techniques and scenarios. This could involve gathering and organizing new datasets that cover various forms of image manipulation. In addition, developing models that can withstand the test of real-world conditions should be conducted using challenging datasets, such as those derived from different camera sources or varying lighting conditions. Furthermore, integrating deep learning frameworks and pre-processing methods can enhance the performance and accuracy of image forgery detectors. In addition, the ethical implications of developing and testing image forgery detection models should be taken into account. This will allow us to ensure that they perform as well as they should in real-world applications.

REFERENCES

[1] E. Aarthi, S. Jana, W. Gracy Theresa, M. Krishnamurthy, A. S. Prakaash, C. Senthilkumar, S. Gopalakrishnan (2022), Detection and Classification of MRI Brain Tumors using S3-DRLSTM Based Deep Learning Model. IJEER 10(3), 597-603. DOI: 10.37391/IJEER.100331.

[2] Bayar, B., & Stamm, M. C. (2016). A deep learning approach to universal image manipulation detection using a new convolutional layer.

International Journal of Electrical and Electronics Research (IJEER)
Research Article | Volume 11, Issue 2 | Pages 631-638 | e-ISSN: 2347-470X

In Proceedings of the 4th ACM Workshop on Information Hiding and Multimedia Security (pp. 5-10).

[3] V Sanjay and P Swarnalatha (2022), Deep Learning Techniques for Early Detection of Alzheimer's disease: A Review. IJEER 10(4), 899-905. DOI: 10.37391/IJEER.100425.

[4] Bayar, B., & Stamm, M. C. (2018). Constrained convolutional neural networks: A new approach towards general purpose image manipulation detection. IEEE Transactions on Information Forensics and Security, 13(11), 2784-2797.

[5] Yang, Y., Luo, Y., & Li, Z. (2019). A novel end-to-end deep learning architecture for image splicing detection. IEEE Transactions on Information Forensics and Security, 14(1), 54-69.

[6] Salloum, S. A., & Bousselham, A. (2018). Image forgery detection using deep neural networks: A survey. Journal of Electronic Imaging, 27(4), 1-16.

[7] Rahmani, H., & Mian, A. (2019). A deep learning approach to universal image manipulation detection using a new convolutional layer. IEEE Transactions on Information Forensics and Security, 14(3), 533-548.

[8] Al-Qershi, O. M., & Ting, T. O. (2019). Deep neural network-based image forgery detection: A review. Artificial Intelligence Review, 51(3), 385-429.

[9] Xu, Y., Wang, S., & Tan, T. (2018). Semi-supervised deep learning for image forgery detection. IEEE Transactions on Information Forensics and Security, 13(9), 2287-2302.

[10] Yang, J., Zhang, J., Wu, W., & Liu, Q. (2019). Image forgery detection based on convolutional neural networks with attention mechanism. In Proceedings of the International Conference on Artificial Intelligence and Security (pp. 46-57).

[11] Qu, C., & Shi, Y. Q. (2018). Universal image forgery detection based on multi-scale convolutional neural network. IEEE Transactions on Information Forensics and Security, 13(10), 2436-2451.

[12] Zhu, J., Lu, H., & Li, H. (2018). Synthetic image forgery detection using GAN and one-class classifier. IEEE Transactions on Information Forensics and Security, 13(11), 2840-2855.

[13] Fridrich, J., 2013. Digital image forensics. IEEE Signal Processing Magazine, 30(3), pp.150-155.

[14] Li, W., Guo, J., Lu, X. and Wei, S., 2018. Image forgery detection using convolutional neural networks and clustering. IEEE Transactions on Multimedia, 20(3), pp.533-543.

[15] Zhao, Y., Ren, J., Wang, Y., Zhang, Y. and Liu, L., 2018. A multi-level approach for image forgery detection using convolutional neural network. IEEE Access, 6, pp.10423-10434.

[16] Li, Z., Xie, Y., Li, X., Wang, J. and Wang, S., 2019. Image forgery detection based on convolutional neural networks and feature pyramid networks. IEEE Access, 7, pp.173228-173240.

[17] Sutthiwan, P., Pintavirooj, C. and Chongstitvatana, P., 2020. Deep learning-based forgery detection in digital images using convolutional neural network. IEEE Access, 8, pp.150406-150418.

[18] Farid, H. and Lyu, S., 2003. Detecting hidden messages using higher-order statistics and support vector machines. In Proceedings of the 9th ACM multimedia conference (pp. 491-498).

[19] Chen, M. and Fridrich, J., 2011. Steganalysis of compressed speech using selected higher-order statistics. IEEE Transactions on Information Forensics and Security, 6(3), pp.783-798.

[20] Qian, Y., Dong, J., Tan, T. and Zhu, Y., 2015. Steganalysis of compressed speech using high-order statistical model and SVM. Journal of Signal Processing Systems, 80(3), pp.323-334.

[21] Li, W., Li, X., Li, Z. and Li, S., 2018. Image forgery detection using multi-scale CNN with feature fusion. IEEE Access, 6, pp.45950-45959.

[22] Bappy, J.H., Paul, M., Roy-Chowdhury, A.K. and Roy-Chowdhury, A., 2018. Exploiting spatial structure for localizing manipulated image regions. IEEE Transactions on Information Forensics and Security, 13(2), pp.477-490.

[23] Zhang, J., Wang, T., Liu, L., & Wang, J. (2019). A Novel Deep Learning Method for Image Forgery Detection Based on Multi-Scale Features. IEEE Access, 7, 20422-20433.

[24] Li, Y., Chang, E. C., & Wang, Y. (2019). Multi-scale dense convolutional neural networks for image forgery detection. IEEE Transactions on Information Forensics and Security, 14(4), 1094-1106.

[25] Li, C., Cao, X., Li, Y., & Wang, S. (2019). Detecting Copy-Move Forgery in Images via Convolutional Neural Networks. IEEE Access, 7, 116546-116554.

[26] Sun, Y., Xue, W., & Zhang, H. (2018). A hybrid deep learning network for image forgery detection. IEEE Access, 6, 20088-20099.

[27] Qu, Z., Liu, X., Chen, C. L., & Zhou, J. (2019). Image forgery detection using a novel convolutional neural network. IEEE Access, 7, 92998-93006.

[28] Yang, Y., Huang, J., & Shen, H. T. (2018). Exposing image splicing with inconsistency in noise level. IEEE Transactions on Information Forensics and Security, 13(1), 68-83.

[29] Zhang, J., Wang, T., Liu, L., & Wang, J. (2019). A novel deep learning method for image forgery detection based on multi-scale features. IEEE Access, 7, 20422-20433.

[30] Lin, X., & Li, X. (2018). Image forgery detection using deep convolutional neural network and attention mechanism. IEEE Access, 6, 33738-33747.