Scientific Research Publishing

Journal of Computer and Communications, 2024, 12, 135-151
https://www.scirp.org/journal/jcc
ISSN Online: 2327-5227
ISSN Print: 2327-5219

Insights into Manipulation: Unveiling Tampered Images Using Modified ELA, Deep Learning, and Explainable AI

Md. Mehedi Hasan[1], Md. Masud Rana[1], Abu Sayed Md. Mostafizur Rahaman[2]

[1]Department of Information and Communication Technology, Bangladesh University of Professionals, Dhaka, Bangladesh
[2]Department of Computer Science and Engineering, Jahangirnagar University, Dhaka, Bangladesh
Email: mehedishuvro12@gmail.com, asmmr@juniv.edu

How to cite this paper: Hasan, Md.M., Rana, Md.M. and Rahaman, A.S.Md.M. (2024) Insights into Manipulation: Unveiling Tampered Images Using Modified ELA, Deep Learning, and Explainable AI. *Journal of Computer and Communications*, **12**, 135-151.
https://doi.org/10.4236/jcc.2024.126009

Received: March 22, 2024
Accepted: June 24, 2024
Published: June 27, 2024

Copyright © 2024 by author(s) and Scientific Research Publishing Inc.
This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).
http://creativecommons.org/licenses/by/4.0/

Open Access

Abstract

Digital image forgery (DIF) is a prevalent issue in the modern age, where malicious actors manipulate images for various purposes, including deception and misinformation. Detecting such forgeries is a critical task for maintaining the integrity of digital content. This thesis explores the use of Modified Error Level Analysis (ELA) in combination with a Convolutional Neural Network (CNN), as well as, Feedforward Neural Network (FNN) model to detect digital image forgeries. Additionally, incorporation of Explainable Artificial Intelligence (XAI) to this research provided insights into the process of decision-making by the models. The study trains and tests the models on the CASIA2 dataset, emphasizing both authentic and forged images. The CNN model is trained and evaluated, and Explainable AI (SHapley Additive exPlanation—SHAP) is incorporated to explain the model's predictions. Similarly, the FNN model is trained and evaluated, and XAI (SHAP) is incorporated to explain the model's predictions. The results obtained from the analysis reveals that the proposed approach using CNN model is most effective in detecting image forgeries and provides valuable explanations for decision interpretability.

Keywords

IFD, DIF, ELA, CNN, FNN, XAI, SHAP, CASIA2.0

1. Introduction

Digital image forgery, also known as image tampering, involves manipulating a digital image to deceive the viewer into believing that the image is authentic. The rapid advancement of digital imaging technologies in recent years has made it

easier than ever to manipulate and alter digital images. The prevalence of image editing software and the effortless dissemination of digital content have amplified the challenge of digital image manipulation in the current digital era [1].

1.1. Overview of the Research

This research proposes a noble approach to digital image forensics, combining traditional ELA techniques with advanced deep learning models like CNNs or FNNs [2] [3]. This combination is further enhanced by XAI to ensure transparency and trust in the detection process [3]. By addressing the need for robust IFD methods, this research holds the potential to safeguard against the harmful consequences of image forgery across various domains [4] like journalism, law enforcement, content verification, and digital forensics by ensuring the authenticity of images and supporting investigations with reliable evidence.

1.2. Literature Review

Detecting digital image forgeries is vital in image forensics, achieved through a blend of ELA and advanced Deep Learning models such as CNNs or traditional neural networks like FNNs, supplemented by Explainable Artificial Intelligence (XAI) techniques. ELA scrutinizes error levels to unveil tampering, while CNNs offer superior capabilities in detecting various forms of image manipulation compared to traditional methods. The integration of ELA with CNNs or FNNs has led to robust and efficient forgery detection systems. Additionally, XAI techniques enhance interpretability, facilitating a deeper understanding of the detection process.

Huang *et al*. [5] pioneered an innovative IFD approach that combines convolutional long short-term memory (ConvLSTM) and CNNs, which demonstrated exceptional effectiveness, reaching a 99.2% accuracy rate on the CASIA image forgery dataset.

Pan *et al*. [6] propose a powerful forgery detection method, combining ELA to pinpoint manipulated image regions and metadata analysis to uncover inconsistencies, effectively detecting diverse forgery techniques like copy-move, splicing, and inpainting.

Chen *et al*. [7] evaluated the ELA-CNN model for image forgery detection, demonstrating its superiority to other advanced methods on a public dataset. This suggests ELA-CNN's effectiveness and potential for accurate forgery detection.

Li *et al*. [8] extensively investigated image forgery methods employing various deep learning architectures to assess their efficacy in identifying forged images.

XAI techniques enhance the transparency and trustworthiness of deep learning models in image forensics by improving their understandability and interpretability [4] [9].

Huang *et al*. [5] introduced a novel approach to detect image forgery through a hybrid model, merging ConvLSTM and CNN. The ConvLSTM captures tem-

poral features from image sequences, while the CNN extracts spatial features from image frames. Their model surpasses existing methods on a public dataset of forged images.

He *et al*. [2] and Huang *et al*. [10] conducted comprehensive surveys on image forgery detection methods utilizing deep learning. Their studies offer insights into the strengths and limitations of different CNN and FNN architectures for detecting image forgery.

Various CNN and FNN architectures, including ResNet, VGGNet, DenseNet, AlexNet, and LeNet-5, have been explored for image classification tasks. CNNs like ResNet, VGGNet, and DenseNet have gained popularity for their effectiveness in capturing spatial features from images, while FNN architectures like AlexNet and LeNet-5 offer versatility across diverse tasks by adjusting parameters. Despite CNNs generally outperforming FNNs in image classification due to their spatial feature extraction capabilities, FNNs remain versatile across different tasks compared to CNNs, which are specifically crafted for image-related tasks.

1.3. Objectives

Despite progress in image forgery detection, there are significant gaps in integrating Error Level Analysis (ELA) with deep learning models like CNNs and FNNs, and in employing Explainable Artificial Intelligence (XAI) in image forensics. This study introduces a noble framework that combines ELA and deep learning models to enhance transparency and interpretability in detecting image forgeries.

The primary objective of this thesis is to develop an image forgery detection framework that combines modified ELA, CNN, FNN, and Explainable AI (SHAP). Specifically, the objectives include:

- To implement modified ELA for image forgery detection.
- To design, train, and validate CNN model and FNN model for image classification.
- To integrate XAI using SHAP to interpret the decisions of the models.
- To evaluate and compare the performance of modified ELA, CNN and FNN along with SHAP in image forgery detection.

2. Model Design and Implementation

Deep convolutional neural network (CNN) models have excelled in image classification tasks, notably surpassing other methods in challenges like the ImageNet challenge [11] [12] [13]. This success stems from training these networks on extensive datasets containing millions of images. Notably, pre-trained weights of various models are accessible through platforms like Keras, serving as an open resource for researchers. This availability facilitates transfer learning, enabling the application of knowledge gained from solving one problem to similar tasks, thereby enhancing the efficiency of subsequent endeavors. The following diagram

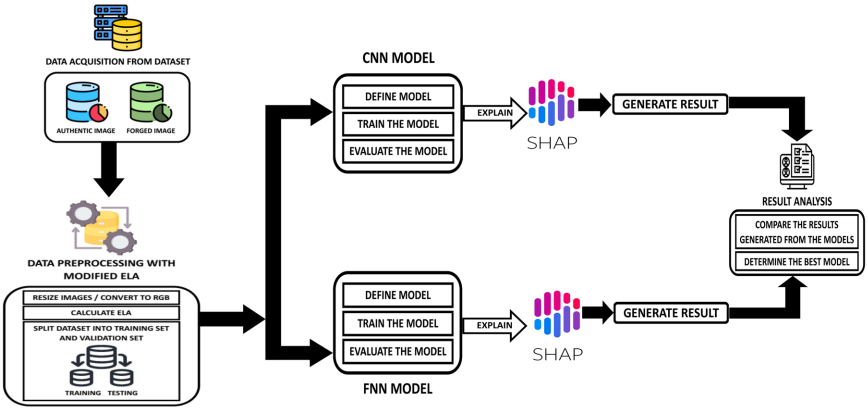

Figure 1. Proposed framework.

(**Figure 1**) illustrated the proposed framework of this research.

The proposed model acquire data from a dataset which contains both authentic and forged images. Subsequently these data are preprocessed with modified ELA algorithm. Preprocessing contains image classification, segmentation and denoising. Through preprocessing, two subsets named as training subset and validation subset are generated. These data are transferred to the models (CNN model and FNN model) for training, validation and evaluation. The results obtained through evaluation; the results are transferred to SHAP explainer. Finally, the results generated from SHAP explainer are compared to decide the best model for detecting image forgeries. The process and relevant topics are described in the subsequent phases of this paper.

2.1. Data Acquisition

The process of collecting and preparing data for analysis is referred as Data acquisition [14] [15]. It involves identifying the sources of the data, collecting the data, cleaning and preprocessing the data, and storing the data in a format that is accessible and usable for analysis [16]. To collect data for this research, the CASIA 2.0 Image Tampering Detection Dataset was used [13].

The methodology entails fine-tuning the model by training the entire network with pre-processed images from the CASIA2.0 dataset [13], which contains both original and tampered images. Standardization involves resizing all images to 224 × 224 pixels and dividing the dataset into training and test subsets. These subsets categorize images as authentic or forged. Preprocessing includes normalization and evaluating compression-induced errors. CNN architecture, aided by ELA [17], discerns authentic from forged images, with XAI using SHapley Additive exPlanations (SHAP) [3] [9] [18] to elucidate predictions. The outcomes of training, testing, and prediction processes determine the superior architecture for image forgery detection.

2.2. Preprocessing with Error Level Analysis (ELA)

Preprocessing, a crucial step in image processing and computer vision applica-

tions like forgery detection, recognition, and classification, involves preparing image data for subsequent processing stages such as feature extraction and analysis, aiming to enhance image quality and extract relevant information. Error Level Analysis (ELA), developed by Neal Krawetz [17], detects image forgery by analyzing errors introduced during image compression, highlighting regions where manipulation likely occurred. Adjustments made during code implementation of the basic algorithm significantly improved forgery detection. The modified algorithm is as follows **Table 1**:

Table 1. Modified Error level analysis.

1	Open the original image.
2	Try to save the original image with the specified quality. If the save is unsuccessful, convert the image to RGB and try to save it again.
3	Open the temporary image.
4	Calculate the absolute difference between corresponding pixels in the original and temporary images.
5	Apply a scaling factor to the absolute difference image.
6	Close the temporary image.
7	Return the scaled absolute difference image.

The flow diagram of the modified algorithm is as follows (**Figure 2**):

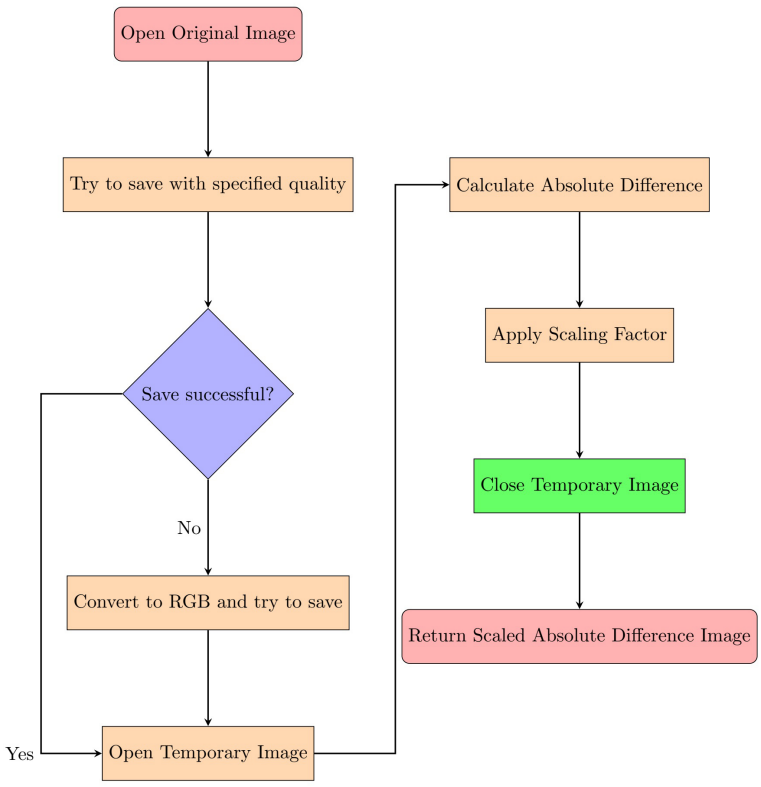

Figure 2. Flow diagram of modified ELA algorithm.

The modified algorithm enhances efficiency and robustness by addressing memory leaks through proper closure of temporary images and eliminating unnecessary loops. Pre-processing with modified Error Level Analysis (ELA) prepares images by emphasizing regions with differing compression levels [19], aiding in the extraction of various image characteristics like texture, edges, and corners. This method enhances the efficiency of tasks such as image classification, segmentation, and denoising.

The major differences between the original algorithm and the modified algorithm are the modified one prevents memory leaks by closing the temporary image and also avoids unnecessary loops. Moreover, the modified ELA algorithm provides more consistency than the traditional one during image handling, which helps maintain a more accurate reference for comparison. Traditional ELA algorithm performs redundant image conversion, which adds more computation time than that of modified ELA algorithm. Thus, the modified implementation becomes more efficient and robust.

2.3. Convolutional Neural Network (CNN)

CNNs, tailored for image and video analysis, mimic human visual processing and excel at identifying complex patterns [12] [20]. They typically consist of convolutional layers for feature extraction, pooling layers for spatial resolution reduction while retaining vital information, and fully connected layers for merging extracted features to produce final outputs like classification labels or segmentation masks [21] [22] [23] [24] [25] (illustrated in **Figure 3**). Convolutional layers slide small filters over input images to extract features, generating feature maps [21]. Pooling layers reduce computational load and prevent overfitting by decreasing feature map resolution [21]. Fully connected layers integrate extracted features to generate final outputs [21].

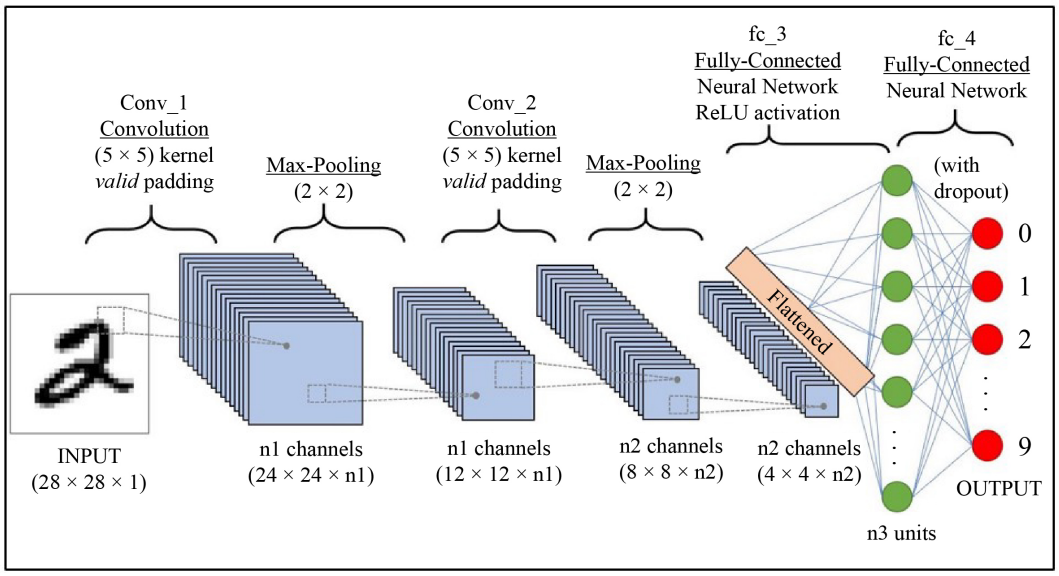

Figure 3. Generic representation of CNN architecture.

The training process of a CNN is governed by several crucial hyperparameters. These include:

- **Learning Rate:** The learning rate dictates the magnitude of adjustments made to the CNN's weights during the optimization process. A higher learning rate accelerates the CNN's learning process but also increases the risk of overfitting the training data.
- **Batch Size:** The batch size is the number of training examples that the CNN updates its parameters on at each step of training. A larger batch size may enhance training efficiency, it can also increase the probability of overfitting the training data.
- **Number of Epochs:** The number of epochs determines how many times the CNN model processes the training data throughout the training phase. More epochs can lead to better performance, but they also take longer to train.
- **Optimizer:** The optimizer is a method that changes the CNN's parameters during training. Adam and SGD are two popular optimizers.
- **Loss Function:** The loss function calculates the error between the predicted output of the CNN and the true output. Two most common loss functions are: Cross-Entropy Loss and Mean-Squared Error.
- **Batch Normalization**: Batch normalization is utilized to stabilize and accelerate the training of the CNN model. It standardizes the inputs of each layer, leading to faster convergence and improved generalization.

2.4. Feedforward Neural Network (FNN)

FNN, an artificial intelligence model composed of interconnected neurons, consists of input, hidden, and output layers [26] [27] [28]. Neurons within an FNN compute weighted sums of inputs and apply activation functions like sigmoid or ReLU to generate outputs. During training, FNN weights are optimized to minimize the loss function, evaluating prediction performance [26]. The FNN model comprises of the Input Layer receives data, while the Hidden Layers process and learn complex relationships, and the Output Layer produces final network output, varying in neuron count based on problem type [29]. A simple FNN model architecture is as follows (**Figure 4**):

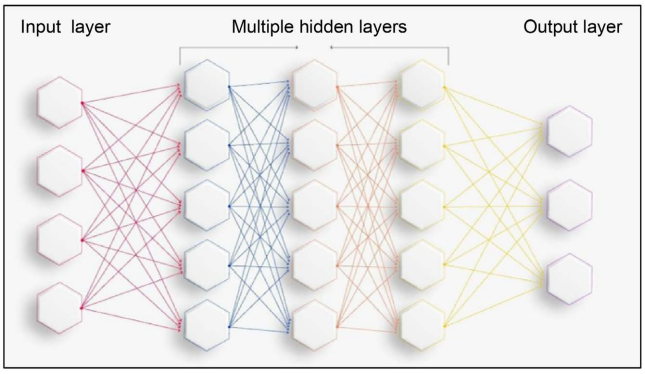

Figure 4. Generic representation of CNN architecture.

FNN hyperparameters regulate the training process, crucial for model performance. Key parameters include:

- **Learning Rate:** The learning rate is a crucial tuning parameter that governs the magnitude of adjustments applied to the model's weights during the training process. A higher learning rate expedites the model's learning but also increases the risk of overfitting the training data.
- **Optimizer:** The optimizer is a method that changes the model's parameters during training, such as its weights and biases. Adam and SGD are two popular optimizers.
- **Number of Epochs:** The epochs represent the frequency with which the model processes the training data throughout the training phase.
- **Number of Layers:** The complexity of the model is determined by the number of layers in an FNN. More complex models can learn more complex relationships in the data, but they are also more likely to overfit.
- **Number of Neurons per Layer:** The capacity of an FNN, or how much complexity it can learn, is determined by the number of neurons in each layer. More neurons mean a greater capacity to learn, but also a greater risk of overfitting.
- **Batch Normalization:** The technique of batch normalization is applied to stabilize and expedite the training process of an FNN model. It works by normalizing the inputs to each layer, which helps the model to converge faster and generalize better.

2.5. Training, Validation, and Evaluation

Training, validation, and evaluation are crucial stages in CNN or FNN model development. Training optimizes model parameters to minimize loss on training data, while validation assesses performance on a held-out dataset to prevent overfitting. Evaluation gauges model performance on a separate dataset for an unbiased estimate of generalization.

To train and evaluate the CNN and FNN model, the CASIA 2.0 dataset is partitioned into two subsets: a training set and a validation set. This division guarantees that the model undergoes training on a specific portion of the data and is then assessed on a separate, unseen portion. The conventional ratio for this partition is frequently established at 80% for training and 20% for validation, ensuring an evaluation of the model's ability to generalize.

The CNN and FNN models' training, validation, and evaluation in this study involve several steps. These include Model Definition (outlining the model structure), Model Compilation (configuring training specifics like loss function and optimizer), Model Training (which trains the model on designated data), and Model Evaluation (assessing its performance on validation data). Callbacks used to enhance training include EarlyStopping (monitoring validation accuracy and halting if stagnant), ReduceLROnPlateau (adjusting learning rate with no validation improvement), and LearningRateScheduler (customizing learning rate

schedule during training).

2.6. Explainable Artificial Intelligence (XAI)

As AI permeates various facets of life, understanding its decision-making processes becomes crucial amid increasing complexity. Explainable AI (XAI), also known as Explainable Machine Learning (XML), addresses this need by making AI models more transparent and comprehensible. XAI offers methods to elucidate how AI models reach decisions, allowing humans to interpret their reasoning. Techniques like Feature Importance, Partial Dependence Plots, Local Interpretable Model-Agnostic Explanations (LIME), and SHapley Additive explanations (SHAP) tailor explanations to different AI models and applications, with SHAP calculating feature impact on predictions, aiding in identifying influential features in the model's predictions.

3. Result and Discussion

The model in this study was implemented using Python 3.12.0 and TensorFlow 2.12.0. Training was conducted on Kaggle Notebook, a cloud-based Jupyter Notebooks environment offered by Kaggle [30], popular among data scientists and machine learning practitioners for its convenience and ease of use. The research evaluated CNN and FNN models, alongside XAI techniques, for digital image forgery detection, showing the proposed model's effectiveness in distinguishing between authentic and forged images.

The proposed CNN and FNN models were evaluated using the CASIA 2.0 Image Tempering Detection Dataset, comprising various forged images subjected to diverse transformations. The dataset was split with 80% for training and 20% for validation and testing. Preprocessing involved modified ELA to emphasize forged regions. The CNN model took about 35 seconds per epoch for training, whereas the FNN model completed each epoch in roughly 4 seconds, demonstrating the suitability of Kaggle notebooks for this research.

The CNN and FNN models, coupled with Explainable AI (XAI) techniques, are evaluated for image forgery detection primarily using accuracy, a fundamental metric reflecting the models' ability to classify images as authentic or tampered accurately. High accuracy signifies effective discrimination between the two categories. Alongside accuracy, other performance metrics such as Precision (indicates low false positive rates), Recall (assesses true positive predictions), F1-Score (measures overall model performance), and Confusion Matrix (visualizes classification accuracy).

3.1. Analysis of Accuracy with Modified ELA Algorithm

Figure 5 and **Figure 6** illustrates the graphs of accuracy (training and validation), as well as the loss (training and validation), for the CNN and FNN models respectively. The CNN model, when combined with modified ELA algorithm, attains a training accuracy of 99.96% and a validation accuracy of 94.21%. On

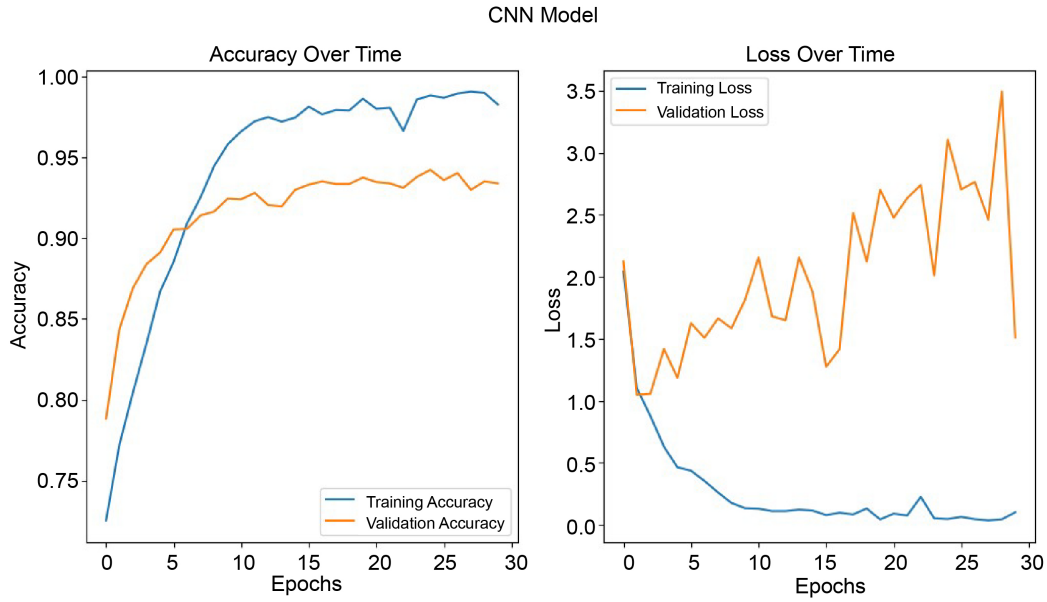

Figure 5. Training and validation accuracy with Modified ELA Algorithm, as well as training and validation loss, of the CNN model (Batch Size: 64 and Number of Epochs: 30).

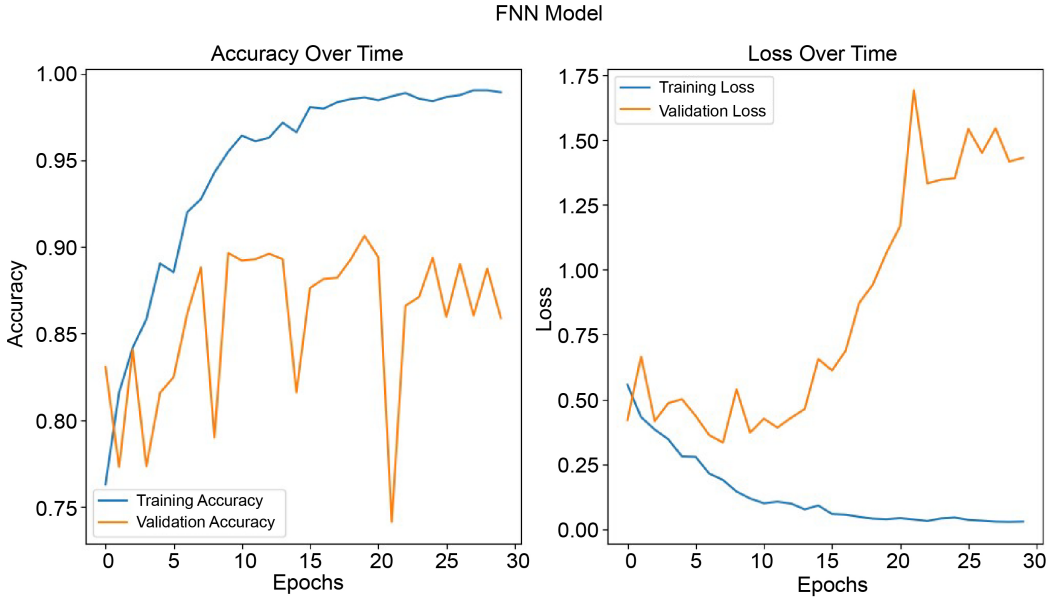

Figure 6. Training and validation accuracy with Modified ELA Algorithm, as well as training and validation loss, of the FNN model (Batch Size: 64 and Number of Epochs: 30).

Table 2. Comparison between CNN and FNN Model with Modified ELA algorithm in terms of accuracy and computational time.

Feature	CNN Model	FNN Model
Training Accuracy	99.96%	99.06%
Validation Accuracy	94.21%	90.65%
Computational Time (Average)	35 seconds	04 seconds

the other hand, the FNN model achieves a training accuracy of 99.06% and a validation accuracy of 90.65%. The CNN model with ELA also achieves a validation loss of 1.5101, while the FNN model achieves a lower validation loss of 1.0654. On contrary, FNN model is computationally faster than CNN model. The summary of the findings in Figure 5 and Figure 6 are presented in Table 2 for better assimilation of the findings.

3.2. Analysis of Accuracy with Traditional ELA Algorithm

Figure 7 and Figure 8 display training and validation accuracy, along with training and validation loss graphs for the CNN and FNN models, respectively. The

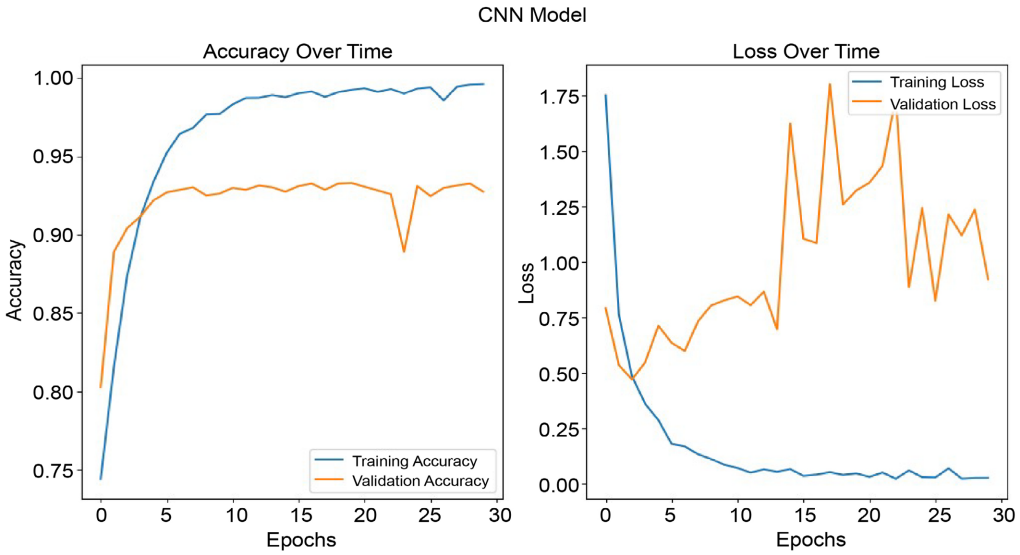

Figure 7. Training and validation accuracy with Traditional ELA Algorithm, as well as training and validation loss, of the CNN model (Batch Size: 64 and Number of Epochs: 30).

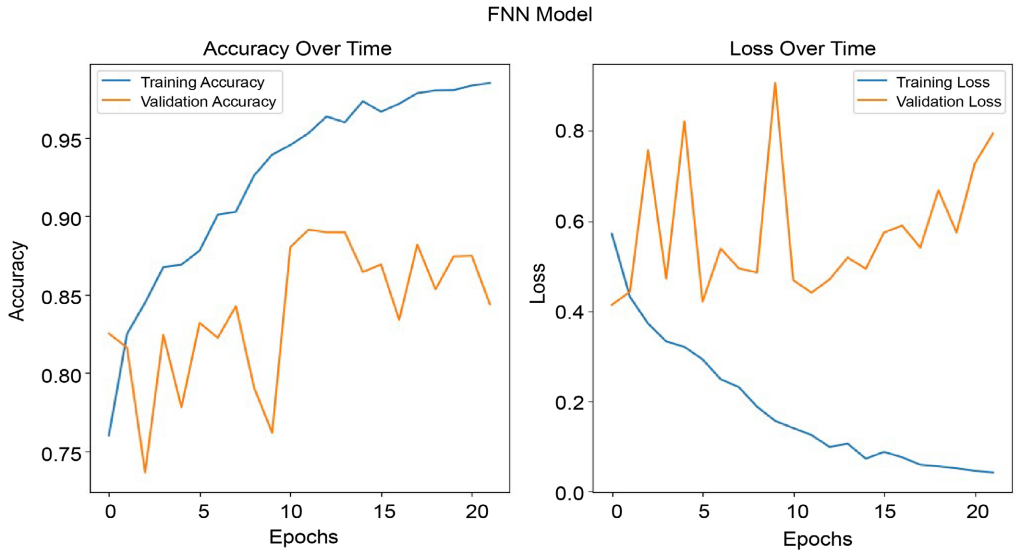

Figure 8. Training and validation accuracy with Traditional ELA Algorithm, as well as training and validation loss, of the FNN model (Batch Size: 64 and Number of Epochs: 30).

Table 3. Comparison between CNN and FNN Model with Traditional ELA Algorithm in terms of accuracy and computational time.

Feature	CNN Model	FNN Model
Training Accuracy	99.62%	98.50%
Validation Accuracy	93.30%	89.14%
Computational Time (Average)	55 seconds	09 seconds

summary of the findings in Figure 7 and Figure 8 are presented in Table 3 for better assimilation of the findings.

With the traditional ELA algorithm, the CNN model achieves a training accuracy of 99.62% and a validation accuracy of 93.30%, with a validation loss of 0.9231. Conversely, the FNN model achieves a training accuracy of 98.50% and a validation accuracy of 89.14%, with a lower validation loss of 0.4411. Despite this, the FNN model exhibits faster computational performance compared to the CNN model.

3.3. Analysis of Additional Performance Metrics

Deep learning models for image forgery detection are evaluated by analyzing performance metrics like precision, recall, and F1-score, illustrated in Figure 9.

Accuracy	Precision	Recall	F1 Score
$\frac{(TP + TN)}{(TP + FN + FP + TN)}$	$\frac{TP}{(TP + FP)}$	$\frac{TP}{(TP + FN)}$	$\frac{2 * (Precision * Recall)}{(Precision + Recall)}$

Figure 9. Metrics definition.

Table 4 presents additional performance metrics for both models utilizing the modified ELA algorithm and the traditional ELA algorithm.

Table 4. Analysis of additional performance metrics with modified ELA algorithm and traditional ELA algorithm.

Model	Precision	Recall	F1-Score
CNN (with Modified ELA Algorithm)	0.94	0.92	0.93
FNN (with Modified ELA Algorithm)	0.91	0.89	0.90
CNN (with Traditional ELA Algorithm)	0.93	0.92	0.92
FNN (with Traditional ELA Algorithm)	0.90	0.88	0.89

3.4. Confusion Matrix

The confusion matrix, a 2 × 2 table, features diagonal elements denoting true positives (TP) and true negatives (TN) for accurate predictions by the classifier, while off-diagonal elements represent prediction errors like false positives (FP) and false negatives (FN). Figure 10 and Figure 11 depict the confusion matrices for the CNN and FNN models with the modified ELA algorithm, respectively.

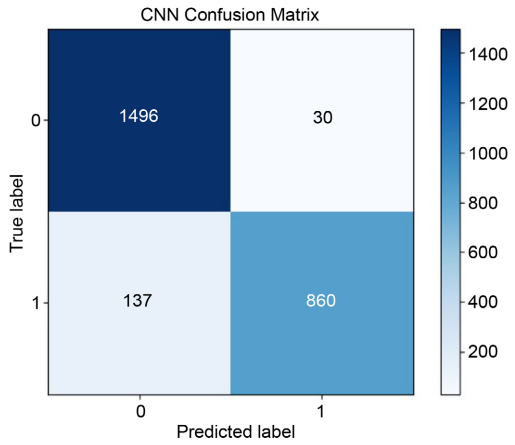

Figure 10. Illustration of the performance of the CNN model with modified ELA algorithm using confusion matrix.

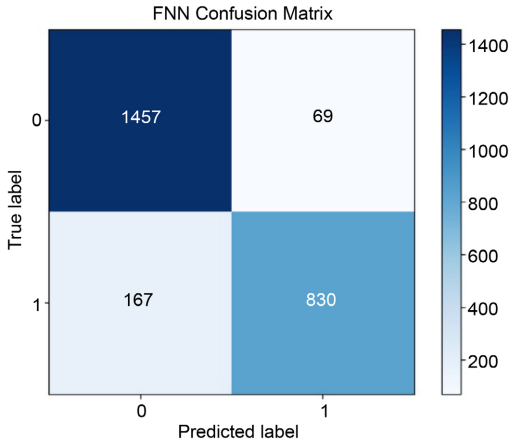

Figure 11. Illustration of the performance of the FNN model with modified ELA algorithm using confusion matrix.

Figure 12 and **Figure 13** display the confusion matrices for the CNN and FNN models with traditional ELA algorithm, respectively.

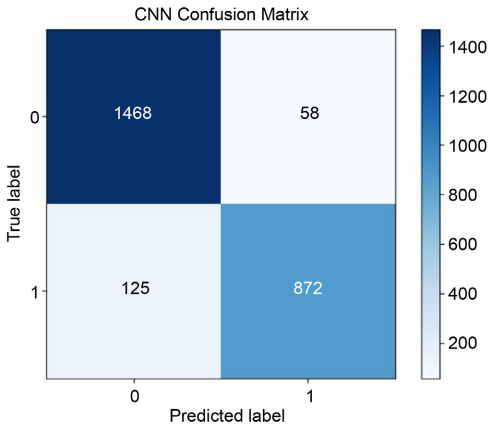

Figure 12. Illustration of the performance of the CNN model with traditional ELA algorithm using confusion matrix.

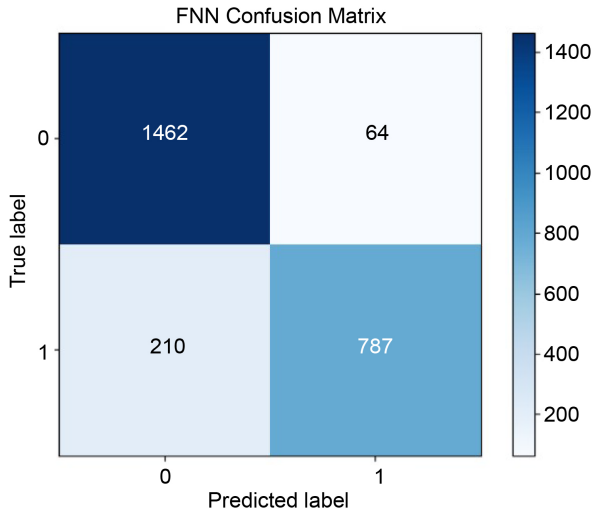

Figure 13. Illustration of the performance of the FNN model with traditional ELA algorithm using confusion matrix.

3.5. Mitigation of the Limitations of Relevant Researches

The model introduced in this study effectively addressed and surpassed the limitations identified in the previously conducted researches addressed in this paper. The proposed model with modified ELA algorithm and CNN model outperformed other models (FNN model with modified ELA algorithm, CNN model with traditional ELA algorithm, and FNN model with traditional ELA algorithm) across all metrics, including training accuracy, validation accuracy, precision, recall, and F1 score. This suggests its superiority in identifying forged images and minimizing false positives and false negatives.

The proposed model could detect all forms of image forgeries with enhanced accuracy. More so, this model could efficiently incorporated SHAP explainer with CNN model and modified ELA algorithm, which could interpret the predictions more precisely. Computational expenses also mitigated by using GPU-based Kaggle kernel.

4. Conclusions

This study utilized deep learning techniques and Explainable AI (XAI) to tackle the challenge of differentiating between genuine and forged images, commonly involving individuals. Image forgery involves malicious alterations to images, often turning authentic images from public platforms into entirely different ones, potentially with inappropriate content aimed at spreading negative publicity. The ELA algorithm plays a crucial role in detecting such manipulation, particularly when input image quality aligns with the algorithm's parameters. The research delved into the efficacy of this combined approach in discerning between authentic and tampered images, yielding several significant findings.

- The CNN model outperformed the FNN model with a test accuracy of 94.21% compared to the FNN model's 90.65%, likely due to CNN's capability in

learning vital spatial features for forgery detection. Integration of XAI further enhances the model's interpretability and transparency in predictions.

- The CNN model outperformed the FNN model across various metrics, including training and test/validation accuracy, precision, recall, and F1-score, indicating its capability in minimizing false positives, detecting forged images, and reducing false negatives.

Overall, the CNN model achieved a test accuracy of 94.21%, which can be comparable to the state-of-the-art methods.

The proposed image forgery detection approach achieved notable success, yet there remains room for improvement. Future endeavors could focus on exploring avenues such as:

- Training and evaluating CNN and FNN models are performed with CASIA 2.0 dataset, as it is widely used and accepted dataset. Future research works may constitute to train and evaluate CNN and FNN models on larger and more diverse balanced datasets of forged and authentic images.
- Developing CNN models that can distinctly detect specific types of image forgeries, such as copy-move forgeries, splicing forgeries, and inpainting forgeries.
- Developing CNN models that can be used to detect forgeries in videos, along with other multimedia content.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

[1] Farid, H. (2009) Image Forgery Detection: A Survey. *IEEE Signal Processing Magazine*, **26**, 16-25. https://doi.org/10.1109/MSP.2008.931079

[2] He, J., Dong, X., Yi, Z. and Wang, H. (2020) Deep Learning for Image Forgery Detection: A Survey. *IEEE Transactions on Information Forensics and Security*, **15**, 3261-3285.

[3] Lundberg, S.M. and Lee, S.-I. (2017) A Unified Approach to Interpreting Model Predictions. *Proceedings of the* 31*st International Conference on Neural Information Processing Systems*, Long Beach, 4-9 December 2017, 4766-4777.

[4] Al-Qazzaz, N. and El-Sawi, M. (2022) Digital Image Forgery Detection Using Deep Learning Approaches: A Comprehensive Review. *IEEE Access*, **10**, 46360-46395.

[5] Huang, H., Huang, W., Zhang, X. and Zhang, Y. (2020) Image Forgery Detection Using Integrated Convolution-LSTM (2D) and Convolution (2D). *IEEE Signal Processing Letters*, **27**, 1561-1565.

[6] Pan, X., Zhang, X., Lyu, S. and Tchoumousset, S. (2010) Image Tampering Detection Using Error Level Analysis and Metadata Analysis. *IEEE Transactions on Information Forensics and Security*, **5**, 492-507.

[7] Chen, J., Luo, W., Li, X., Li, B. and He, J. (2019) ELA-CNN: A Hybrid Model for Digital Image Forgery Detection. *IEEE Transactions on Information Forensics and Security*, **14**, 150-162.

[8] Li, Y., Li, X. and Li, B. (2020) Comprehensive Study on Image Forgery Techniques Using Deep Learning. *IEEE Transactions on Information Forensics and Security*, **15**, 3261-3285.

[9] Ribeiro, M.T., Singh, S. and Guestrin, C. (2016) Why Should I Trust You?: Explaining the Predictions of Any Classifier. *Proceedings of the* 22*nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, San Francisco, 13-17 August 2016, 1135-1144. https://doi.org/10.1145/2939672.2939778

[10] Huang, H., Huang, W., Tan, T. and Li, J. (2017) A Novel Passive Image Forgery Detection Algorithm Based on Block-Level Error Level Analysis. *IEEE Transactions on Information Forensics and Security*, **13**, 60-74.

[11] He, K., Xiangyu, Z., Shaoqing, R. and Jian, S. (2016) Deep Residual Learning for Image Recognition. *Proceedings of* 2016 *IEEE Conference on Computer Vision and Pattern Recognition*, Las Vegas, 27-30 June 2016, 770-778. https://doi.org/10.1109/CVPR.2016.90

[12] Simonyan, K. and Zisserman, A. (2014) Very Deep Convolutional Networks for Large-Scale Image Recognition. *Proceedings of* 2014 *IEEE International Conference on System Engineering and Technology*, Bandung, 24-25 November 2014, 354-355.

[13] Dong, J., Wand, W. and Tan, T. (2013) CASIA Image Tampering Detection Evaluation Database. 2013 *IEEE China Summit and International Conference on Signal and Information Processing*, Beijing, 6-10 July 2013, 422-426. https://doi.org/10.1109/ChinaSIP.2013.6625374

[14] Jain, A.K. (2010) Data Clustering: 50 Years beyond K-Means. *Pattern Recognition Letters*, **31**, 651-666. https://doi.org/10.1016/j.patrec.2009.09.011

[15] Witten, I.H., Frank, E. and Hall, M.A. (2011) Data Mining: Practical Machine Learning Tools and Techniques. Morgan Kaufmann.

[16] Mahalanobis, P.C. (1936) On the Generalised Distance in Statistics. *Proceedings of the National Academy of Sciences of India*, **2**, 49-55.

[17] Krawetz, N. (2016) Error Level Analysis (ELA) for Image Forensics. *Digital Forensics Investigation*, **19**, 16-28.

[18] Li, Y., Li, X. and Li, B. (2021) Explainable Image Forgery Detection Using SHAP. *Proceedings of the* 2021 *IEEE International Conference on Image Processing* (*ICIP*), Anchorage, 19-22 September 2021, 1401-1405.

[19] Gonzalez, R.C. and Woods, R.E. (2018) Digital Image Processing. 4th Edition, Pearson Education.

[20] LeCun, Y., Bottou, L., Bengio, Y. and Haffner, P. (1998) Gradient-Based Learning Applied to Document Recognition. *Proceedings of the IEEE*, **86**, 2278-2324. https://doi.org/10.1109/5.726791

[21] Guo, Y., Liu, Y., Krizhevsky, A. and Fei-Fei, L. (2019) Understanding Convolutional Neural Networks with a Mathematical Model. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, **49**, 506-520.

[22] Sharma, P. (2023) Basic Introduction to Convolutional Neural Network in Deep Learning. https://www.analyticsvidhya.com/blog/2022/03/basic-introduction-to-convolutional-neural-network-in-deep-learning/

[23] Rosebrock, A. (2023) Convolutional Neural Networks (CNNs) and Layer Types. https://pyimagesearch.com/2021/05/14/convolutional-neural-networks-cnns-and-layer-types/

[24] GeeksforGeeks (2023) CNN: Introduction to Pooling Layer.

https://www.geeksforgeeks.org/cnn-introduction-to-pooling-layer/

[25] Amidi, A. and Amidi, S. (2023) CS 230-Convolutional Neural Networks Cheat-Sheet. https://stanford.edu/~shervine/teaching/cs-230/cheatsheet-convolutional-neural-networks

[26] Goodfellow, I., Bengio, Y. and Courville, A. (2016) Deep Learning. MIT Press.

[27] Hagan, M.T. and Menhaj, M.B. (1994) Training Feedforward Networks with the Marquardt Algorithm. *IEEE Transactions on Neural Networks*, **5**, 989-993. https://doi.org/10.1109/72.329697

[28] Turing (2023) Understanding Feed Forward Neural Networks with Maths and Statistics. https://www.turing.com/kb/mathematical-formulation-of-feed-forward-neural-network

[29] Shivkumar, D. (2023) Introduction to Feed Forward Neural Network. https://www.scaler.com/topics/deep-learning/introduction-to-feed-forward-neural-network/

[30] Kaggle (2023) Kaggle Notebook. https://www.kaggle.com/docs/notebooks