International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

An Improved Deep Learning-Based Image Steganalysis

A M Yola[1], A. Y. Gital[2], Usman Ali A[2], K. M. Musa[3], Hajara Musa[3]

[1]ICT Directorate, Federal University of Kashere Gombe, Nigeria

[2, 3] Department of Mathematical Sciences, Abubakar Tafawa Balewa University Bauchi, Nigeria

[2]Department of Computer Sciences, Federal collage of Education Technical Gombe, Nigeria

[3]Department of Computer Sciences, Gombe State University, Gombe, Nigeria

ABSTRACT

Steganalysis involves identifying and examining covert data concealed through steganography in digital content. Nevertheless, current activation functions like ELU, ReLU, and Leaky ReLU have drawbacks, including dead neurons and high computational demands. A novel hybrid activation function that integrates a convolutional neural network (CNN) to streamline the feature extraction and classification processes has been suggested to solve the weakness associated with negative values. According to the regression results, the proposed approach surpasses the existing activation functions, with a mean square error of 6%.

1.0 Introduction

Steganalysis is a technique to uncover hidden messages in digital media like images, sounds, videos, and plain-text documents(Guo et al., 2014; Sahu & Sahu, 2020). This involves hiding messages within a "Transporter/Cover media" to establish secret communication(Fatnassi et al., 2016; Ranjan & Forensics, 2016). Steganography uses different algorithms such as specific, target, blind, or universal and hiding techniques like Steghide, OpenStack, COVER, TEST, JUNIWARD, and UERD to hide messages in images(Yan et al., 2021; Yola et al., 2023). In image steganalysis, machine learning and deep learning techniques such as CNN, RNN, and DNN are used to overcome embedding algorithms(Beke & Kumbasar, 2019; Sabeena & Abraham, 2021). The proposed approach uses CNN as the modelling process with hybridized activation functions to demonstrate an advancement over existing works. The system simplifies feature extraction and classification procedures while potentially improving their effectiveness by combining them into a cohesive end-to-end framework(Yu et al., 2020; Zoph & Le, 2018). However, distinguishing image content with different texture intricacies can be challenging due to the variances produced by steganographic indicators(D. Hu et al., 2021; J. Hu et al., 2019; Li et al., 2014). To address this, the evaluation highlights feature-based steganalysis approaches that select and absorb relevant information from cover images(Gustavo et al., 2019; Rathika et al., 2017). The study also offers insights into advanced image steganography and its five types: spatial, transform, spread spectrum, statistical, and distortion(Atta & Ghanbari, 2021; Eid et al., 2022; Z. Wang et al., 2020). Thus, the research shows how the regression result is presented using training and evaluation datasets.

However, the review notes that the performance of current CNN-established approaches is still not as robust as eccentric handcrafted feature-based approaches(Yedroudj et al., 2018). To improve detection accuracy, the study proposes different methods. Firstly,handcrafted feature-based steganalysis and CNN-based techniques enhance detection accuracy. Finally, the review offers a model that utilizes three hybridized activation functions to improve the performance of the GNCNN architecture(Hussain et al., 2020). Overall, the research provides an insightful analysis of steganalysis, steganography delving into the different types and approaches used to uncover hidden image messages.

2.0 Literature Review

Neural Network The domain of neural networks is an attractive realm to explore. These intricate systems are crafted to imitate biological networks, acquiring the ability to execute tasks using an introduction to diverse datasets and patterns, avoiding the dependence on task-specific rules. They are predicated on models of computational threshold logic, deriving inspiration from either cerebral investigations or the application of neural networks to the realm of artificial intelligence(X. Hu et al., 2021; Montavon et al., 2018; X. Wang et al., 2022; Z. Wang et al., 2020). Moreover, neural networks have contributed to the advancement of the finite automata concept(Gupta et al., 2021; Scardapane et al., 2019; Q. Wu & Wu, 2022). A classical neural network encompasses neurons, connections, weights, biases, propagation functions, and a learning rule(Zhang et al., 2019). The process of learning in neural networks entails the adjustment of free parameters, such as weights and biases, which can modify the network's responsiveness to its surroundings(Bellaby, 2021; Gital et al., 2022; X. Hu et al., 2021). Similarly, the study highlights also the evolution of deep learning.

Deep learning mirrors the evolution of neural networks(F. Hu et al., 2021). Starting in the 1950s, advancements in computer hardware technology paved the way for the transformation of neural networks from single-layer models to multi-layered structures, ultimately giving rise to the contemporary deep neural networks we are familiar with today(Bach et al., 2015). Neural networks, according to Hecht Nielsen, an American neural network scientist, a neural network is a computing system made up of several simple, highly interconnected processing elements, which process information by their dynamic state response to external inputs(Hecht-Nielsen, 1992). Moreso, neural networks are made up of an input layer, one or more hidden layers, and an output layer. Each neuron connection has a weight that is modified during training to improve the network's performance(Kato et al., 2020).

In addition, deep learning techniques have grown in recent years as a result of increased processing power and data generation(D. Hu et al., 2021). Therefore, deep learning helps to produce robust AI systems thatsimply could not be possible some years ago(Erickson et al., 2017). Thus, Large-scale data interpretation and information generation are made quicker and simpler using deep learning(L. Wang et al., 2022; Zeng et al., 2019). It is used in a wide range of fields, including automatic driving, medical equipment and image processing(Cao, 2022) that focuses on how deep learning providesa solution to embedding mystery information on images using deep learning techniques(Okajima & Sadamasa, 2019). The biggest advantage of using deep learning in research is the ability to increase layers andtrain the model to execute feature engineering by itself(Aggarwal, 2018; Tabares-Soto et al., 2019)

2.1 Analysis of Existing Algorithms

This study aimed to analyze the algorithms employed in digital image Steganalysis from 2014 to 2021. The review examined these algorithms' advantages, limitations, and incentives. Rectified linear activation (ReLU) emerged as a popular motivational function, while quantization was the least utilized activation function(Chhikara et al., 2018). For further details on the algorithms employed in digital images by scholars during this period, please refer to Table 1 below. Furthermore, this study delved into the benefits and drawbacks of the Steganalysis Algorithm in the digital image domain, with a specific focus on jpeg and spatial domains.

Table 1: ReLU activation function

Reference and Year	**Algorithm**	**Activation Function**
Rectified Linear Units are used as an activation function.		
Fatnassi, A., Gharsellaoui, H., &Bouamama, S. (2016). *A New Hybrid Steganalysis Based Approach for Embedding Image in Audio and Image Cover Media.*	Pitfalls-Net2016	ReLU
G. (2017). *Deep convolutional neural network to detect J-UNIWARD.*	Xu-Net V3 2018	ReLU
Yang, J., Shi, Y.-Q., Wong, E. K., & Kang, X. (2017). *JPEG Steganalysis Based on DenseNet.*	Yang-Net 2018	ReLU
Boroumand, M., Chen, M., & Fridrich, J. (2019). *Deep residual network for steganalysis of digital images.*	SR-Net 2018	ReLU
Yedroudj, M., Comby, F., & Chaumont, M. (2018). Yedroudj-net: *An efficient CNN for spatial steganalysis.*	Yedroudj-Net2018	ReLU
Zhong, S., & Chen, K. (2018). *A novel steganalysis method with deep learning for different texture complexity images.*	Zhong-Net 2019	ReLU
Zeng, J., Tan, S., Liu, G., Li, B., & Huang, J. (2019). WISERNet: *Wider Separate-Then-Reunion Network for Steganalysis of Color Images.*	WISER-Net2019	ReLU

Table 1:Thereview shows how Pitfalls-Net 2016, Xu-Net version 3, 2018, Yang-Net 2018, and SR-Net 2018 used ReLU. Similarly, Yedroudj-Net 2018, Zhong-Net 2019, and WISER-Net 2019 all use the ReLU activation function. The Pitfall-Net shows how experiments are carried out in the clairvoyant scenario. Then, using the fully connected neural network (FNN) allows the learning process to run through a single global optimization. In Xu-Net V3, the model uses a single channel of information in the selective statistic,thereby showing the importance of Data Augmentation in image steganalysis.Furthermore, Yang-Net uses CNN architecture and prediction in the Jpeg domain for their experiment. Also, SR-Net usedminimized and visible force elements that showed excellent performance. Each of these mentioned models uses both spatial and Jpeg domains. The drawback of Yang-Net is that the execution is more profound and slower.SR-Net has the weakness of the condensed nature of connections, and performance was not satisfactory.

Moreover, Yedroudj-Net 2018, Zhong-Net 2019, and WISER-Net 2019 all approach factor ReLU as an activation function. The model of Yedroudj-Net uses pre-processing with the 30-filter bank. The model uses truncation activation and batch normalization with a linked scale layer. Yedrouj-Net performed better than Ye-Net and Xu-Net. The shortcoming ofYedroudj-Net's model is its complexity, leading to a time increase. Classical Zhong-Net trained in image texture feature and grey level co-occurrence matrix. The datasets appear divided as subsets with many intricacies. Training took place in the most effective region (MER). The disadvantage of this method seems to be a weakness of the adaptive algorithm in network generalization. More so, the Wiser-Net 2019 model revealed steganalysis structure in JPEG color images. The model introduces Channel-wise convolution that makes the network performance superior. The model's limitation isits complexity, with several mismatches in the process. The chart below, Figure1, shows a diagrammatical representation of the activation function ReLU.

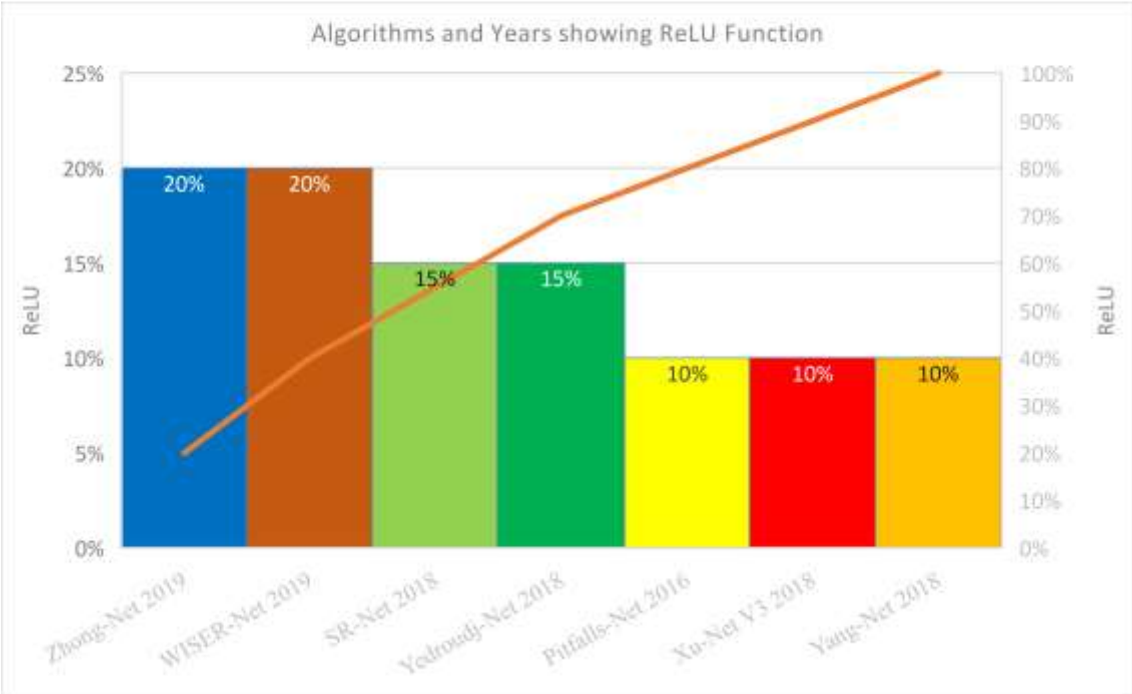

Figure 1: Graph Showing ReLU activation function used by different Algorithm

Table 2: Two activation functions

Reference and Year	**Algorithm**	**Activation Function**
Rectified Linear Units are used as an activation function.		
Y.Qian, J. Dong, W. Wang, (2015). *Deep learning for steganalysis via convolutional neural networks.*	Qian-Net 2015	ReLU & TanH
G. Xu, H.-Z. Wu, (2016). *Structural design of convolutional neural networks for steganalysis*.	Xu-Net v1 2016	ReLU &TanH
J. Ye, J. Ni, & Y. Yi (2017). *Deep learning hierarchical representations for image steganalysis*.	Ye-Net 2017	ReLU & TLU
M. Chen, V. Sedighi, M. Boroumand, & J. Fridrich (2017). *Jpeg-phase-aware convolutional neural network for steganalysis of jpeg images*.	Chen-Net 2017	ReLU and TanH
You, W., Zhang, H., & Zhao, X. (2021). *A Siamese CNN for Image Steganalysis*.	You-Net 2020	ReLU & TLU

Table2, the review shows how Algorithms such as Qian-Net 2015, Xu-Net Version 1 2016, Ye-Net 2017, Chan-Net 2017, and You-Net 2020were discussed with their benefit and shortcomings. The Qian-Net model was unified into a single architecture called Gaussian-Neuron CNN (GNCNN). Likewise, the Xu-Net was found to be connected with enhanced noise residual, low-cut filter, and high pass filter layer network. Then, the Yet-Net and Chan-Net show benefits in data argumentation and information through a single-channel modification of the CNN architecture that improves detection accuracy. The shortcoming of the Qian-Net model is the non-convergence of network nodes. The model's performanceis less superior to the Spatial Rich Model (SRM), making it expensive to run on high resolution. The limitations of Xu-Net are low performance at the early stage and lack of proper application of CNN. Since data sets are Auto-encoder enabled in Xu-Net. Also, limitations of Yet-Net are slow speed andensemble methods are not generic. The Chan-net limitation is a time of training in polynomial neural networks that appears longer and more complex. Similarly, You-Net (SiaStegNet)shows CNN-based architecture consisting of two symmetrical subnets with shared parameters in three phases: pre-processing, feature extraction, and classification. The main benefit of the You-Net is in the unification of the network layer and the heterogeneous nature of the data sets. The shortcoming of the You-Net Model is that the model has yet to solve arbitrary-size image steganalysis. There are also many mismatches in the statistical distribution caused by the resizing of images. Figure2, below,representsthe activation function ReLU with TLU or TanH in a chart.

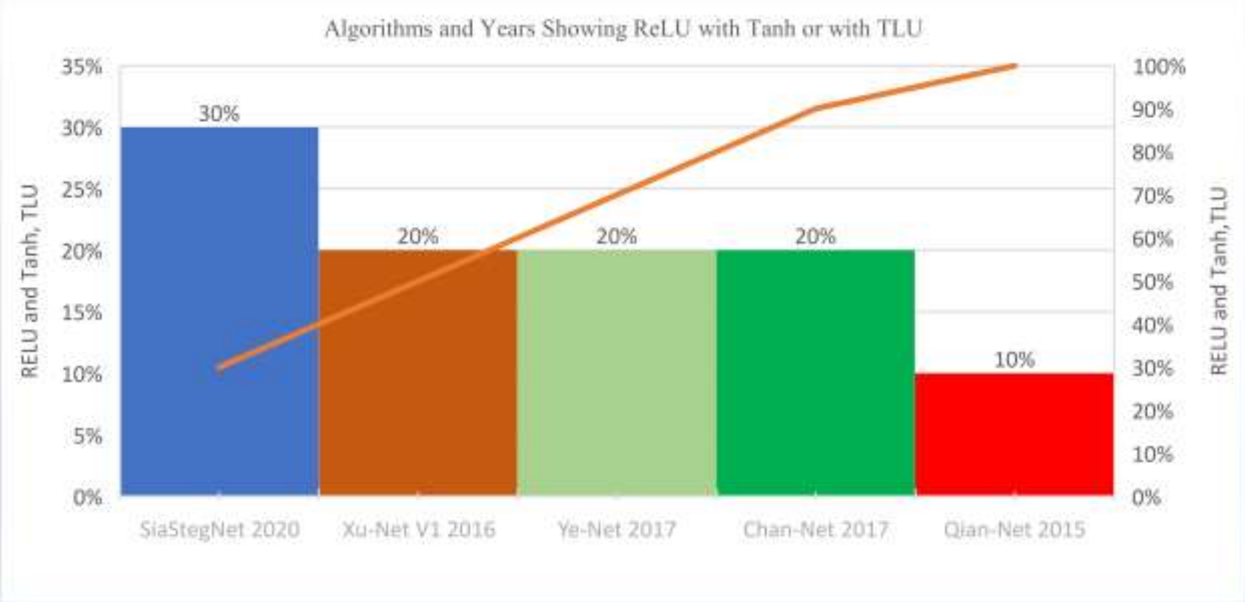

Figure2: Graph Showing Activation Function ReLU used with either TLU or TanH

Table 3: Two Different Activation Functions

Reference and Year	**Algorithm**	**Activation Function**
Rectified Linear Units and Other activation functions		
Li, B., Wei, W., Ferreira, A., & Tan, S. (2018). ReST-Net: *Diverse Activation Modules and Parallel Subnets-Based CNN for Spatial Image Steganalysis*	ReSt-Net 2018	ReLU and Logistic function (Sigmoid)
Qian, Y., Dong, J., Wang, W., & Tan, T. (2017).*Feature learning for steganalysis using convolutional neural networks*.	Qian-Net 2017	Logistic Regression (Sigmoid)
Zeng, J., Tan, S., Liu, G., Li, B., & Huang, J. (2019). WISERNet: *Wider Separate-Then-Reunion Network for Steganalysis of Color Images.*	Zeng-Net 2018	Quantization and Truncation
Zeng, X., Feng, G., & Zhang, X. (2019). Detection of double JPEG compression using a modified DenseNet model.	Zeng-Net V2 2018	BN and ReLU
Wu, S., Zhong, S., & Liu, Y. (2018). *Deep residual learning for image steganalysis.*	Wu-Net2018	Gaussian Activation and ReLU
Tan, S., Wu, W., Shao, Z., Li, Q., Li, B., & Huang, J. (2021). CALPA-NET: *Channel-Pruning-Assisted Deep Residual Network for Steganalysis of Digital Images.*	Calpa-Net 2020	Truncated Linear Unit (TLU)

Similarly, from Table 3, the review shows ReSt-Net 2018, Qi-Net 2018, and Zeng-Net version II 2018 factor the use of non-linear filters in some instances. Then, diverse activation modules and parallel subnets-based CNN for spatial image steganalysisperformbetter than Xu-Net and Ye-Net without Spatial and Channel-wise Attention (SCA). The limitation of the approach was the slowness of the model, and the structure was vast. The qi-Net model allows subjective size or resolution for the input. Meanwhile, the system consists of several convolutional layers that are fully connected. The Structureshows a design based on a Multi-column Convolutional Neural Network (MCNN). The methoduses an activation step and a pooling step. The performance of Qi-Net is better than that of Ensemble Classifier with SRM features, as reported. The disadvantage of the model is having a small set of steganography algorithms; likewise, there is information loss due to downsampling. The Zeng-Net version II model adopted densely connected convolutional networks known as Dense-Net. Zeng-Net achieves complete detection through JPEG compression involving thoroughly combined and pooling layers. Likewise, dense block, single-compression, and image density are experienced in the model. The limitation of Zeng-Net is in training the datasets using dropout techniques.

More so, Wu-Net 2018 and Calpa-Net 2020 hinted at the development of image steganalysis. For instance, Wu-Net simplifies a Deep Residual Learning Network (DRN). The model indicated the value network layer in many digital image samples. DRN preserves the stego signal important to the Insight of the cover and stego image in the process. The Wu-Net Model performs better than the classical rich model method. The Wu-Net is expensive in computational resources and thus lacks efficiency. Likewise, the detection error rate increases, which helps in overfitting occurrences in the approach. Furthermore, Calpa-Net reveals a hybrid criterion making use of three stages. The target is that it does protect datasets. Meanwhile, Calpa-Net used three steps: training, pruning, and fine-tuning. The stages ensure the protection of sets of inherited learned essential weights. The model improved the manner of adaptiveness, transferability, and scalability while the convolutional layer appeared in a data-driven way. Likewise, the limitation of the model includes performancethat is slightly worse than the original SR-Net. When tested with a Jpeg stego image. Meanwhile, the resulting redundancy parameters make the model slower. Figure4 shows the diagrammatical representation.

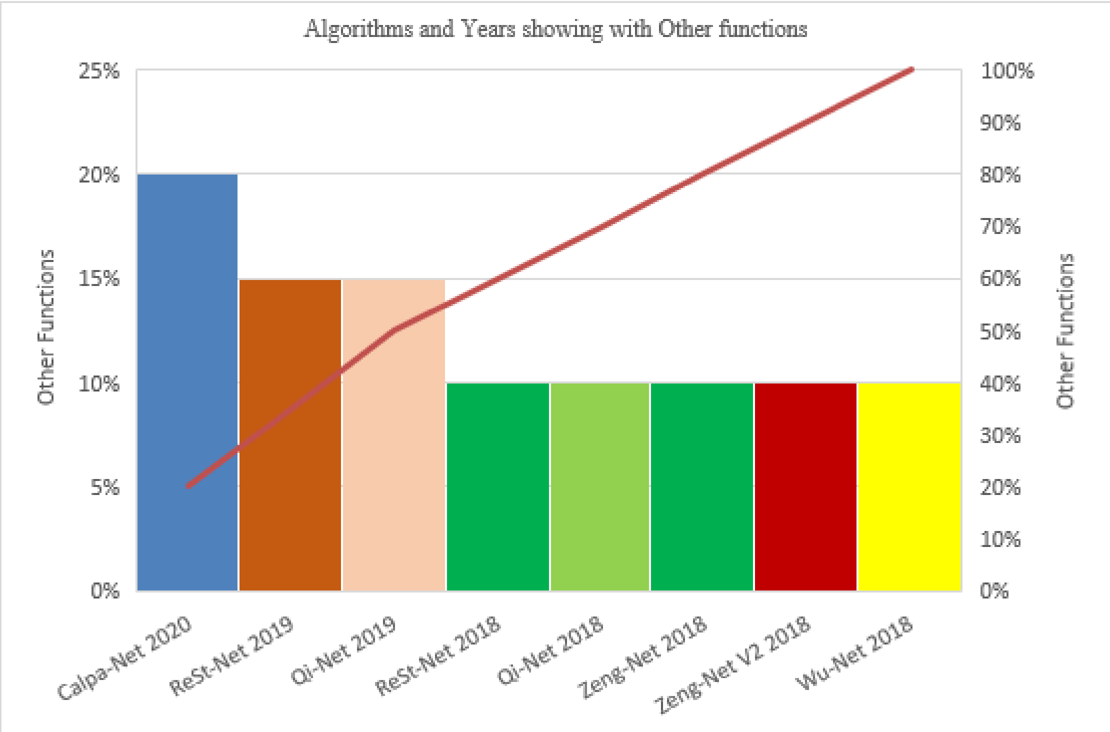

Figure 4: Graph Showing Activation Functions with Other Functions in different years

The review results indicate that several techniques require improving deep learning concepts in image steganalysis. The results also suggest that the Truncated Linear Model (TLU) significantly decreases in effect, leading to a loss in performance. However, the Gaussian process in convolution reduces the error in learning the model. The depth in the 152 layer has an error rate of 3.57% on the test set. Meanwhile, the Residual Network had a single-model error of 4.49% when tested on 152 layers. The complexity result shows a lower error percentage of 3.57% when experimented. However, the model shows fewer filters with significant baseline FLOPS. Furthermore, most of the model uses CNN in structuring Digital image steganalysis, with the Spatial-rich model experiment showing a fair result. Most of the datasets contain 10,000 units of grayscale cover images. However, errors were detected after training at 0.48, 0.43, and 0.31 using an auto-encoder-based blind step analyzer. Meanwhile, SPAM and SRM stood at 0.42 and 0.14, respectively, using the deep learning toolbox for the experiment. Likewise, the handcrafted steganalysis feature and deep learning show improvement and effectiveness based on some perspectives. The optimal simulator is for data embedding. The simulator's accuracy stood at 73.21% at 0.4 bpnz, in which 10,000 grayscale images with 512x512 sizes are considered from BossBase. The optimal simulator is for data embedding. In addition, a Deep Residual learning-based Network (DRN) encompasses many network layers used to capture the digital images' compound statistics, but overfitting occurrence becomes truthful and expensive(S. Wu et al., 2018). Also, the result of the Most Effective Region (MER) approach used in training shows effectiveness with different texture complexities, showing 90% improvement compared to other models like HUGO. More so, detection accuracies of exercise show that the algorithm models on additional texture complexity images are 82.1% and 92.6% in WOW and 81.4% and 90.0% in S-UNIWARD. In most cases, the resultant models highlighted the need to improve the speed of the simulations. For instance, the rate of the training model shows that ordinary texture images are faster than complex texture images. The training takes 40 to 50 epochs of iteration to unite the network for regular texture. Forharshsurfaces, training takes about 70 to 80 generations of iteration. Furthermore, authors like Zhong-Net show that results are built on the idea of MER. Several datasets are distributed into subsets with distinctive intricacy based on the index steganalysis, and the network is trained using the same pattern and inception ideas. This highlights the need for further improvement in the components like complexity, time, and accuracy in image steganalysis. In a similar development, some approaches look asymptotic; for instance, the Gaussian distribution allows for drawing a closed-form expression of control for the robust detectors in the selection. While the Neyman–Pearson optimality criterion is used for a given false alarm probability, I have reviewed the text you provided and corrected any spelling and grammar errors. Please find the revised text below:

The review results indicate that several techniques require improving deep learning concepts in image steganalysis. The results also suggest that the Truncated Linear Model (TLU) significantly decreases in effect, leading to a loss in performance. However, the Gaussian process in convolution reduces the error in learning the model. The depth in the 152 layer has an error rate of 3.57% on the test set. Meanwhile, the Residual Network had a single-model error of 4.49% when tested on 152 layers. The complexity result shows a lower error percentage of 3.57% when experimented. However, the model shows fewer filters with significant baseline FLOPS. Furthermore, most of the model uses CNN in structuring Digital image steganalysis, with the Spatial-rich model experiment showing a fair result. Most of the datasets contain 10,000 units of grayscale cover images. However, errors were detected after

training at 0.48, 0.43, and 0.31 using an auto-encoder-based blind step analyzer. Meanwhile, SPAM and SRM stood at 0.42 and 0.14, respectively, using the deep learning toolbox for the experiment. Likewise, the handcrafted steganalysis feature and deep learning show improvement and effectiveness based on some perspectives. The optimal simulator is for data embedding. The simulator's accuracy stood at 73.21% at 0.4 bpnz, in which 10,000 grayscale images with 512x512 sizes are considered from BossBase. The optimal simulator is for data embedding. In addition, a Deep Residual learning-based Network (DRN) encompasses many network layers used to capture the digital images' compound statistics, but overfitting occurrence becomes truthful and expensive. Also, the result of the Most Effective Region (MER) approach used in training shows effectiveness with different texture complexities, showing 90% improvement compared to other models like HUGO. More so, detection accuracies of training show that the algorithm models on different texture complexity images are 82.1% and 92.6% in WOW and 81.4% and 90.0% in S-UNIWARD. In most cases, the resultant models highlighted the need to improve the speed of the simulations. For instance, the speed of the training model shows that ordinary texture images are faster than complex texture images. The training takes 40 to 50 epochs of iteration to unite the network for ordinary texture. Meanwhile, for complex textures, training takes about 70 to 80 epochs of iteration. Furthermore, authors like Zhong-Net show that results are built on the idea of MER. Several datasets are distributed into subsets with distinctive intricacy based on the index steganalysis, and the network is trained using the same pattern and inception ideas. This highlights the need for further improvement in the components like complexity, time, and accuracy in image steganalysis. of $\alpha 0 = P(\alpha(x) = H1|H0)$ in arriving at the findings. Warden (2018) introduced how the detection affected utilization when Warden did not exploit the knowledge of the selection channel. The result shows the cover image in an 8-bit grayscale with pixel value $Z = (Z1\ldots, ZN)\ Zn \in (0\ldots, 255)$. The percentage of accuracy is above normal standards when compared to other algorithms like HUGO that are shown in the experiments that are carried out on BOSS base on 1.01 containing 10,000 grayscales 512×512 images that are achieved with FLD ensemble permit instead of binary outputs. In most of the models, activation functions such as Rectified Linear Activation (ReLU) and Hyperbolic Tangent (TanH), while in some instances, Logistic function (Sigmoid) and Logistic Regression. The best performance of the size filter in the spatial domain is seven instead of 3, which shows a significant improvement by at least 1.10%. Therefore, the spatial and JPEG domain results show 2% and 3% achievement, and a slight reduction of 0.24% in security performance was experienced. The analysis also shows that the segment of all of the wrap/stego images is categorized into sizes in a small circumspect model. The B×B (B=8b, b=2, 3…, min (M, N/8) where each chunk is divided B2/ 64 DCT block of size, the larger the size of testing, the more precise the output. Although the computational is high, in a situation where the sample is randomly selected in the testing from Sample K, the selection. The classification is based on the grey level at 8×8 size and 64 grey level values and derived steganalysis characteristics. The maximum performance of the Gaussian distribution method over the penny method is almost close to 15% with an embedding rate at 0.2bpd, and in using overlapping in training 400 images, the percentage increases from 71.82% to 80.96%, 81.72%, 82.06%, and 82.39% respectively. Therefore, the above findings answered the research question on the current trends of steganalysis algorithms,motivation function, benefits, and challenges within the investigation period, i.e., 2014-2021. The algorithms and motivational functions are used as the correlation characteristics utilized in the reviewed papers to stimulate quick learning in the field of image steganalysis, and that answers the second research question. Moreover, the review also discovered that most of the reviewed research papers focused on hidden layers using activation functions such as ReLu, which also answered research question three in this study.

3.0 Proposed Framework

A problem must exist before designing a model, and the dataset used is crucial for improving complexity(Gliner et al., 2021). The algorithm is for classification and regression but cannot be directly functional or extended to deep models(Zhong & Chen, 2018). Likewise, the architecture of the deep learning process is at the bottom, the unifying and automated approach is shown as a single-step approach in Figure5:

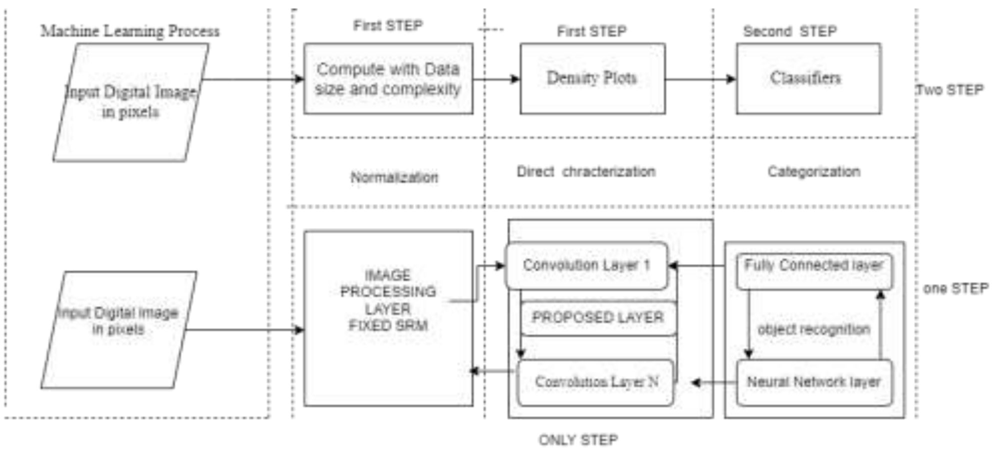

Figure5: The Proposed Framework.

Figure 5 shows the proposed improved complexity of image steganalysis; theoffered techniques/ approaches to use deep learning-based feature extraction are Convolutional Neural Networks (CNNs) - CNNs are a popular choice for image classification and regression tasks(You et al., 2021). For the steganalysis aspect, it is proposed that CNNs are trained on a stego and cover images dataset to extract relevant features for distinguishing between the two(Ke et al., 2019). In steganalysis, transfer learning can be used to fine-tune a pre-trained CNN on a dataset of stego and cover images(Shorten &

Khoshgoftaar, 2019). In the research work, the acquired knowledge is implemented on the dataset and used in a collaborator environment to run the pre-training test. How the model concept works in the hidden layer

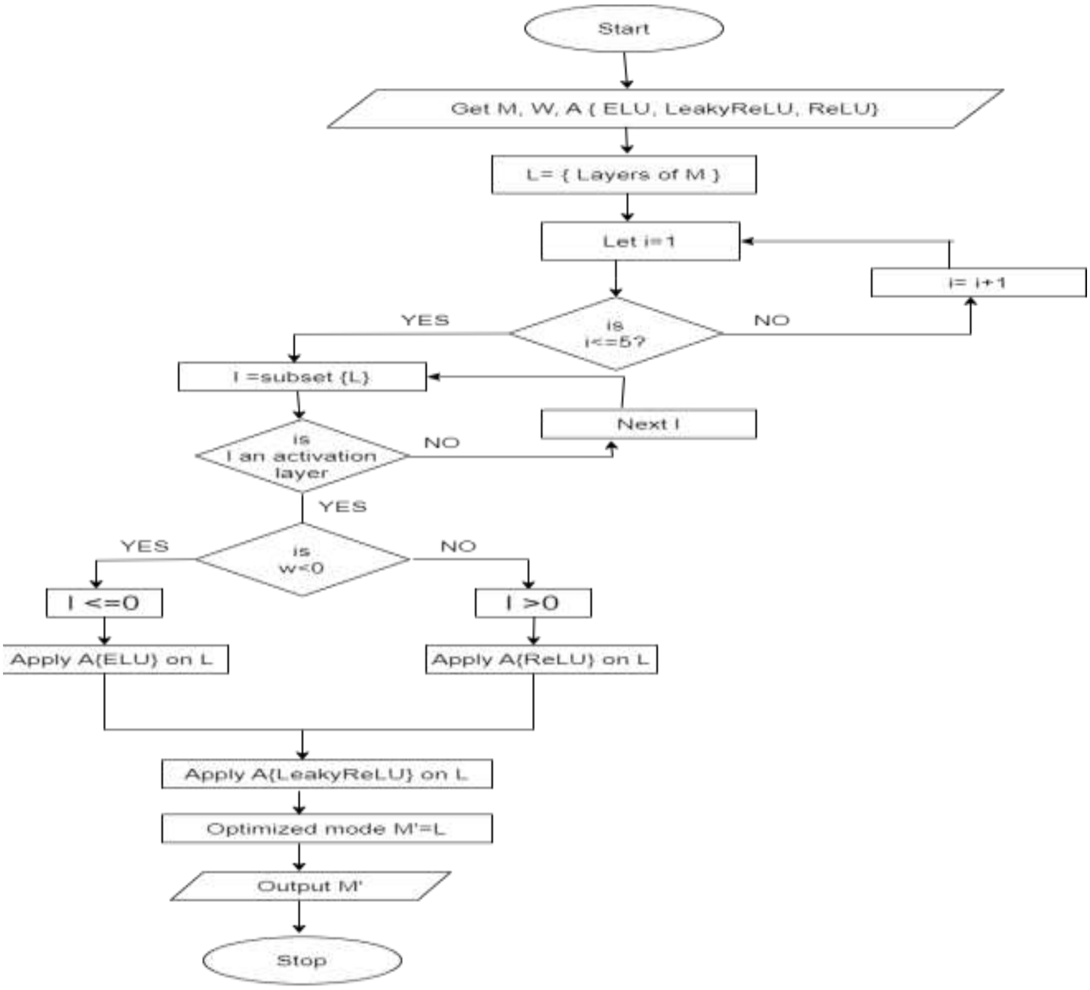

Figure6: Flowchart of the proposed model.

3.1 Expended Model Hidden Layer

The figure 6 displayed in the expanded model concept shows the cycle of the hybridized activation function, which provides a promising value of the digital image after running a test. Initially, all alfa values in the table are set to default, while the filters keep changing or increasing from 64 to 128 from layer 1 to layer 5 in the hidden layer. In deep learning, a filter refers to the weights or parameters of a convolutional layer in a neural network. It is used to process input data and create a feature map displayinginput features, such as edges, shapes, colors, or patterns in a digital image. The filter size is the size of the filter, such as 3x3 or 5x5, and the effectiveness and efficiency of a convolutional layer are influenced by size and testbusing the Adam optimizer, which rates model parameters in real-time using adaptive learningrates and moment estimation to improve the accuracy and speed of CNNs. Thus, have some similarities and differences, such as

$$\boldsymbol{ELU}(\boldsymbol{x}) = \begin{cases} \boldsymbol{if}\ x \geq 0 \\ \propto (\boldsymbol{e}^{x} - \mathbf{1})\ \boldsymbol{if}\ x < 0 \end{cases} \text{-------------------} \ \mathbf{(1)}$$

ELU activation Function

where α is a hyperparameter regulating the value, ELU approaches negative inputs.

The proposed stimulation function arranged average activations arereduced bias sand grade. curve enhances learning speed accuracy, it crucial for digital imagecomplexity,defineas

$\boldsymbol{ReL}(\boldsymbol{x}) = \mathbf{max}(\mathbf{0}, \boldsymbol{x})$-----on is definedas bold as bas

$$\boldsymbol{Leaky\ ReLU}(\boldsymbol{x}) = \begin{cases} x\ if\ \ x \geq \mathbf{0} \\ \boldsymbol{a}x\ \ if\ x < 0 \end{cases} \text{------------------------}(3)$$

LeakyReLU function with α=0.1:

Equation 3 is a leaky ReLU function showing how leaky ReLU is defined; the

Adding an activation function to ReLU can be slower due to the exponential operation and can cause neurons to become inactive if inputs are negative. The Explicit Linear Unit (ELU) is hybridized with the two activation functions, addressing negative inputs to improve outcomes.

Traditional learning processes have limitations, and deep learning methodologies have specific hyperparameters that enhance performance and efficiency. These hyperparameters include learning rate, epoch, dropout rate, max pooling, filters, and hidden layers. These improvements reduce manual effort, improve fitting capacity, prevent overfitting, and enhance the model's robustness. An additional epoch, rising to 30 to 50 generations, is considered to optimize our model better. The proposed kernel increase in CNNs aims to improve training and validation by balancing accuracy and efficiency. A larger kernel size increases the number of parameters and receptive field but also increases computational cost and overfitting risk(Yuan et al., 2019). The kernel moves on input data with a stride value, downsampling the input to reduce overfitting (Lin et al., 2020). Batch normalization is introduced to make information more consistent, speeding up the training process and reducing overfitting. This technique is used before activation functions. Another improvement to the hybridized model is in max pooling, a type of pooling technique that involves taking the supreme value within each rectangular area of the feature map. Furthermore, in the model training, max pooling was used after the convolutional layer in CNNs (3,3). The highest value in each pixel window is shrinking, making the features suitable and lowering the computational cost. In addition, the pooling was specified, giving transparency over the input, while the window moved the stride each time. As a result, a max pooling layer's output has fewer stimulations than the input and highlights the most significant feature within each region. The impression exploited the efficiency and the ability to decrease the spatial dimensionality of the feature maps. Thus, making the CNN architecture more resistant to small disparities in the input digital image. This study uses dropout to reduce overfitting in neural networks (NN) by ignoring some nodes or units in a layer during training. The dropout rate p represents the chance of a neuron being dropped out in each training step. Con2D convolution operation applies kernel size and shape to the input data using a fully connected layer.

3.2 Dataset

The datasets in the research are divided into (2), which include the training and the validation or testing. Meanwhile, the proposed research uses 80% of a dataset for training, with 20% for testing. The model uses splitting, scaling, and normalization to ensure better performance. The data is divided into training and testing to prevent overfitting and optimize the model(Yedroudj et al., 2018). Normalization is applied for proper optimization. The model's performance is tested using various activation functions, and features are transformed to a similar scale to prevent data leakage. The research uses accuracy, precision, F1 score, and Racall to measure the model's effectiveness. The TensorFlow library, an open-source software for building and deploying ML and deep learning models, is used to compute data flow graphs.

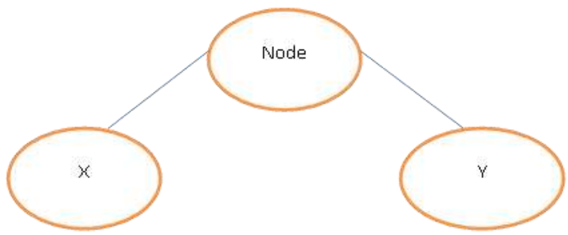

Figure 7: Typical Example of a data flow

An example of data flow is shown in Figure7: Data flow is essential in the hidden layer of the activation function because it determines how the information from the input layer is transformed and propagated to the output layer. Thus, the distribution and handling of different operations or nodes in a neural network is called data flow(Zhou et al., 2020). The node in the DFD is a data store, an external entity, or a process. Data flows that enter or exit a node have labels X and Y(Yang et al., 2017). For example, X is the input data name for the process node, and Y is the output data name for the same node. X and Y can also be variables that represent the data values that are manipulated or stored by the nodes. Keras is a favored wrapper for TF(Ravi & Alazab, 2021). The choice of environment is the Google Collaboratory

3.3 Hybridized Algorithm

The main reason for the hybridization of the model is the activation function; below is the table of other algorithms, benefits, shortcomings, and motivation functions. They show different algorithm methodologies for achieving the model goal.

Table 4:

Algorithms and Year	**Benefits**	**Shortcomings**	**Motivation Function**
Yedroudj-Net 2018	The model uses pre-processing about 30-filter bank, truncation activation, and five convolutional layers with Batch Normalization linked using a Scale Layer. The model is better than Ye-Net and Xu-Net.	Complexity of the mode, convergence time increases	ReLU
Zhong-Net 2019	The model uses the Image texture feature, a gray-level co-occurrence matrix. Several data sets are divided into subsets of different intricacy based on the index steganalysis, developed and trained in the MER-Most Effective Region and initial ideas.	Network generalization capability on the adaptive algorithm is weak	ReLU
WISER-Net 2019	Steganalysis structure for JPEG color images. Channel-wise convolution was introduced to the network, which makes performance superior.	The model is very complex, with mismatching	ReLU
You-Net (SiaStegNet) 2020	Siamese, CNN-based architecture, which consists of two symmetrical subnets with shared parameters, and contains three phases: preprocessing, feature extraction, and fusion/classification. Unfixed network layers. Datasets are heterogeneous	Methodology yet to solve arbitrary-size image steganalysis, Mismatch in the statistical distribution caused by resizing of images.	ReLU and TanH
CALPA-NET 2020	The hybrid criterion usesthree stages: training, pruning, and fine-tuning to protect a set of inherited learned necessary weights. The convolutional layer is in a data-driven manner. The model shows adaptivity, transferability, and scalability.	CALPA- SRNet is slightly worse than the original SRNet when JPEG stego images are with Q F 95. This indicates that when used to detect 0.4 bpp, Redundant parameters make it slower.	Truncated Linear Unit (TLU)

Table 1shows different models from 2018 and 2020. Showing some activation functions ReLU and TLU, ReLU and Tanh, and ReLU only. Therefore, this research study presents a precise, distinct method of hybridizing different activation functions. Where the hybridized algorithm gives the step-by-step of the newly proposed model. Thus, the CNN architecture for the model was designed by hybridizing different activation functions with varying parameters, kernel values, max-pooling layers, and dropout layers.

The research work shows the hybridized algorithm model has shown promising results.

The hybridized model is shown as follows:

```
Function HybridizeLUActivationFunctionInCNNForSteganalysis
Input:
        Let M be the model.
        Let W be the input-weighted vectors in the model.
        Let A {ELU, LeakyReLU, ReLU} be the set of activation functions.
Output:
        A modified model M'
Algorithm:
        Let L be the set of layers of M.
        For i = 1 to 5
            For each l in L
                If l is an activation layer, Then
                    If W < 0 Then
                        Select l' from A {ELU}
                     Else, If W > 0, Then
                        Select l' from A {ReLU}
```

End If

End If

If l' is still <= 0 Then

Select l' from A {LeakyReLU}

End If

End For each

Replace l with l' in L.

End For

Create a new, improved model M' from L.

End

4.0 RESULTS

4.1 In this section, the research presents results and analysis showing the three algorithms used to improve complexity in image steganalysis. The result improves the problem of vanishing gradient and dying neurons. The following parameters are used in our training and testing:They are

The parameters setting during the training phase and testing phase of the proposed approach

Table 5:

Hyper-parameter	values
Learning rate	0.0001
Maximum Epoch	50
Dropout rate	0.2
Batch- size	32
Validation sample	0.20
Training sample	0.80
Kernel size	(5,5)
Max pooling	Five by 5

4.4 REGRESSION ANALYSIS OF DIFFERENT MODELS

4.4.1 Mean Square Error Graph of the ELU activation function

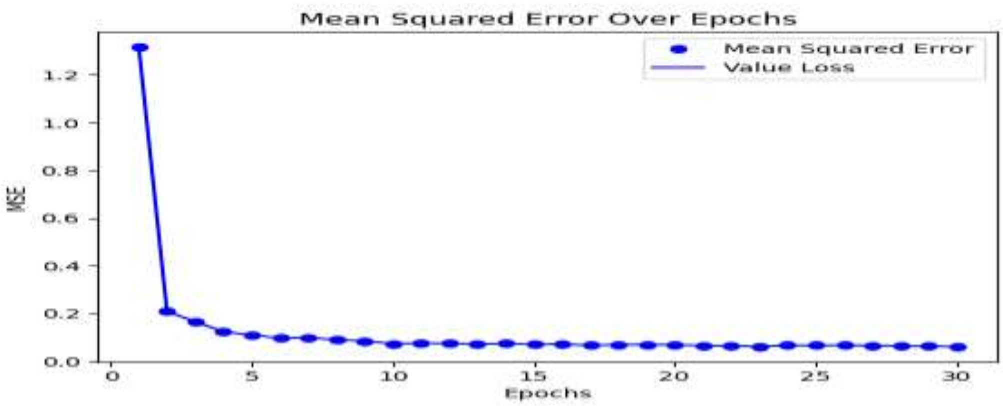

Figure 8: The ELU activation function error over epoch

The graph in Figure 8 demonstrates a notable reduction in the mean square error rate across the epoch. This reduction can be attributed to utilizing the ELU activation function, which can compute negative values. Specifically, the value of 0.08 indicates a superior performance in handling training datasets with lower error rates.

4.4.2 Mean Square Error Graph of the ReLU Activation Function

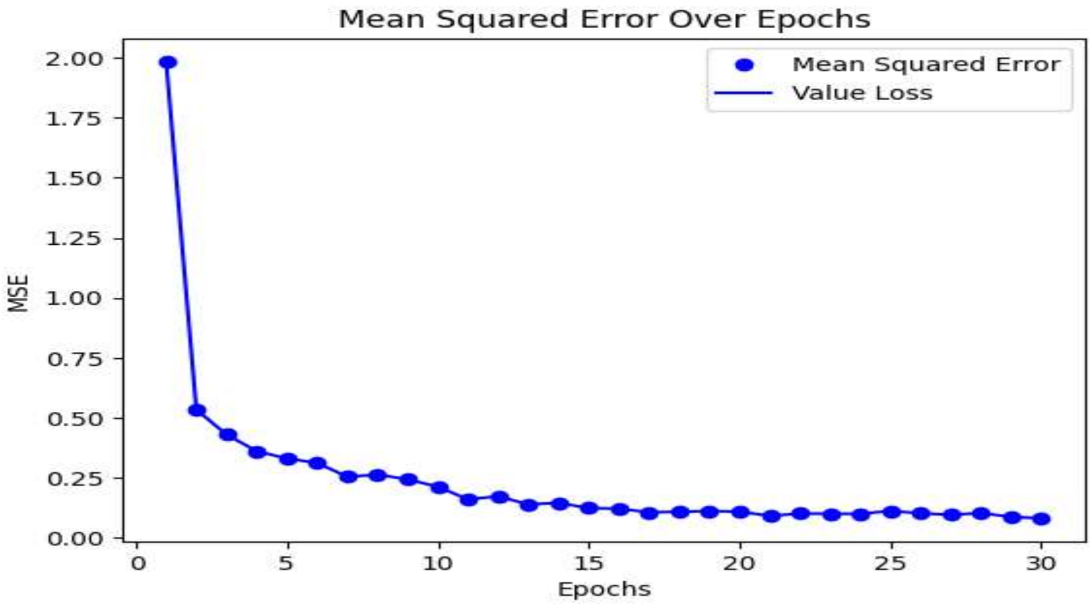

Figure 9: Mean Square Error Graph of the ReLU activation function

The graph displayed in Figure 9 represents the ReLU activation function. It shows the error rate at the 30 iterations and the value loss. Although the performance is relatively good, there is some concern about dying or inactive neurons that ReLU cannot handle, which produced a value of 0.125.

4.4.3 Mean Square Error Graph of the LeakyReLU Activation Function

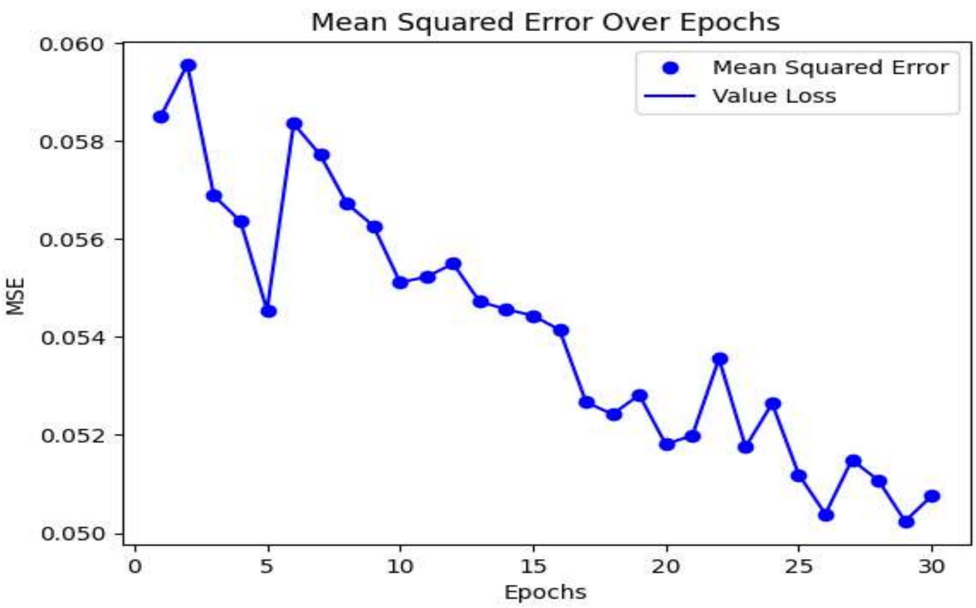

Figure 10: The LeakyReLU activation Function MSE over epoch

Figure 10shows that the LeakyReLU activation function performed better than ReLU because of its effectiveness in handling small negative slopes; thisindicates that the performance is better when compared to

4.4.4 Mean Square Error Graph of the CNN +HAF model.

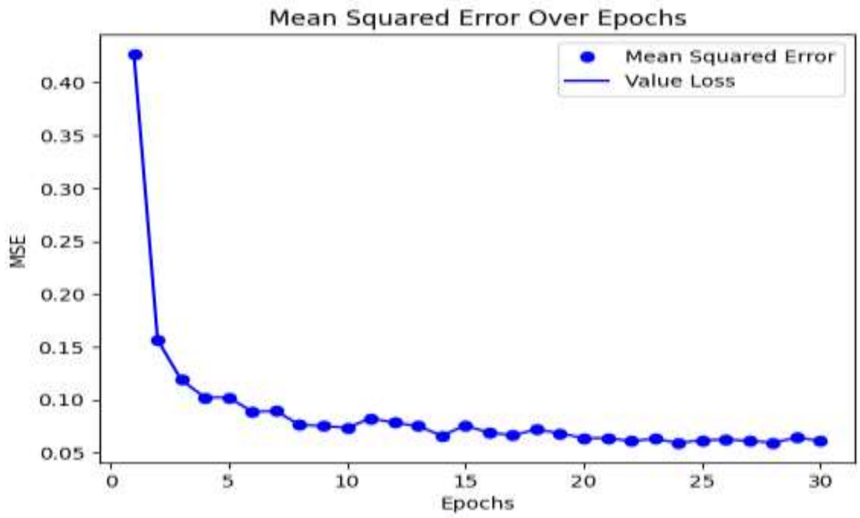

Figure 11: Mean Square Error Graph of the hybridized CNN +HAF model.

In Figure 11, the graph shows almost 0.6 after just 30 epochs in the training; a period is a single pass of the training dataset, allowing the model to learn and improve its performance. In addition, MSE and ages are crucial in training CNN. MSE measures the difference between predicted and actual values and is used to minimize loss. Monitor changes in MSE between epochs to balance underfitting and overfitting. Therefore, the hybridized activation function shows better improvement when compared to existing models.

4.5 REGRESSIONAL ANALYSIS OF THE PROPOSED MODEL AGAINST THE EXISTING MODELS

In deep learning regression, performance evaluation is vital. The analysis resulted from finding a more robust way of improving the existing metrics based on Mean Squared Error (MSE), which calculates the average squared difference between predicted and actual values.

Table 6: Values of MSE for the Activation Functions

Activation Function	**Mean Square Error**
ELU	0.0800
ReLU	0.1250
Leaky ReLU	0.0508
Hybridised (ELU+ReLU+Leaky Relu)	0.060

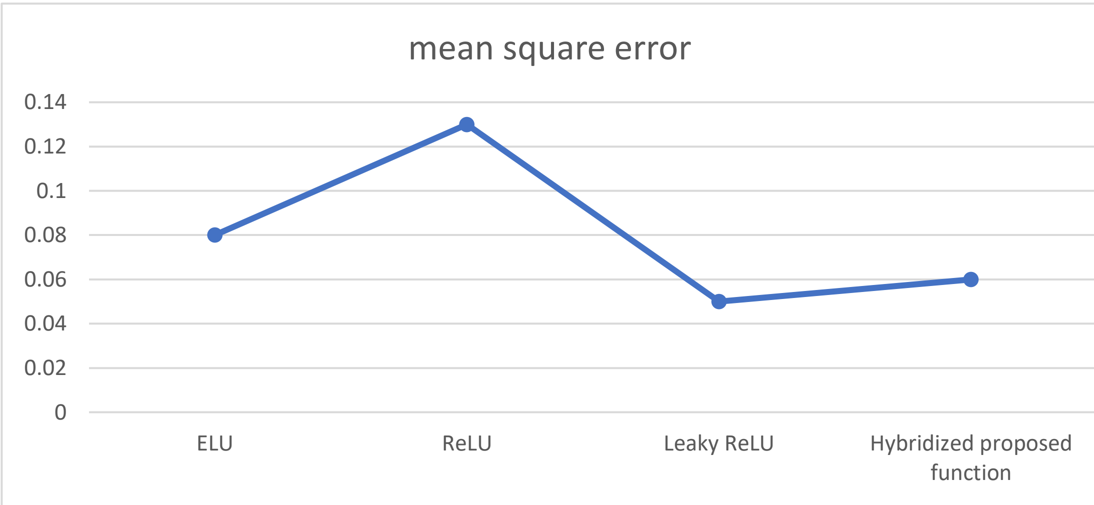

Figure 12: Mean Square Error Chat

In Table 6, the value of the hybridized activation function in CNN is lower, showing 0.06, as opposed to ELU and ReLU, whose values are 0.08 and 0.125, respectively. Therefore, the mean square error of the hybrid activation function shows improvement compared to the existing models. LeakyReLU mitigates the "dying ReLU" problem, resulting in better training performance.

Figure 12 displays a graphical representation of the model's performance, where CNN+HAF demonstrates better results with less error in both the training and evaluation datasets. This indicates that the hybridized activation function performs better than the existing models.

REFERENCE

Aggarwal, C. C. (2018). Neural Networks and Deep Learning. In Neural Networks and Deep Learning. https://doi.org/10.1007/978-3-319-94463-0

Atta, R., & Ghanbari, M. (2021). A high payload data hiding scheme based on dual tree complex wavelet transform. Optik, 226, 165786. https://doi.org/10.1016/j.ijleo.2020.165786

Bach, S., Binder, A., Montavon, G., Klauschen, F., Müller, K. R., & Samek, W. (2015). On pixel-wise explanations for non-linear classifier decisions by layer-wise relevance propagation. PLoS ONE, 10(7), 1–46. https://doi.org/10.1371/journal.pone.0130140

Beke, A., & Kumbasar, T. (2019). Learning with Type-2 Fuzzy activation functions to improve the performance of Deep Neural Networks. Engineering Applications of Artificial Intelligence, 85, 372–384. https://doi.org/https://doi.org/10.1016/j.engappai.2019.06.016

Bellaby, R. W. (2021). An Ethical Framework for Hacking Operations. Ethical Theory and Moral Practice, 24(1), 231–255. https://doi.org/10.1007/s10677-021-10166-8

Cao, L. (2022). Deep Learning Applications. IEEE Intelligent Systems, 37(03), 3–5. https://doi.org/10.1109/MIS.2022.3184260

Chhikara, R. R., Sharma, P., & Singh, L. (2018). An improved dynamic discrete firefly algorithm for blind image steganalysis. International Journal of Machine Learning and Cybernetics, 9(5), 821–835. https://doi.org/10.1007/s13042-016-0610-3

Eid, W. M., Alotaibi, S. S., Alqahtani, H. M., & Saleh, S. Q. (2022). Digital Image Steganalysis: Current Methodologies and Future Challenges. IEEE Access, 4. https://doi.org/10.1109/ACCESS.2022.3202905

Erickson, B. J., Korfiatis, P., Akkus, Z., Kline, T., & Philbrick, K. (2017). Toolkits and Libraries for Deep Learning. J Digital Imaging, 30, 400–405. https://doi.org/10.1007/s10278-017-9965-6

Fatnassi, A., Gharsellaoui, H., & Bouamama, S. (2016). A Novel Approach for Embedding Information in Image Cover Media with Security. Proceedings - 2016 World Symposium on Computer Applications and Research, WSCAR 2016, 10(6), 104–109. https://doi.org/10.1109/WSCAR.2016.27

Gital, A. Y., A, U. A., Kwami, A. M., Zambuk, F. U., & Aminu, A. A. (2022). GAX and GAS Modes : A New Hybrid Deep Learning Method for Phishing URL Detection using GRU , Attention Mechanisms , SVM and XGBOOST Algorithms. 10(3), 426–436.

Gliner, J. A., Morgan, G. A., Leech, N. L., Gliner, J. A., & Morgan, G. A. (2021). Measurement Reliability and Validity. Research Methods in Applied Settings, 319–338. https://doi.org/10.4324/9781410605337-29

Guo, L., Ni, J., & Shi, Y. Q. (2014). Uniform embedding for efficient JPEG steganography. IEEE Transactions on Information Forensics and Security, 9(5), 814–825. https://doi.org/10.1109/TIFS.2014.2312817

Gupta, A., Gupta, A., Narwaria, R. P., Narwaria, R. P., & Prof, M. S. (2021). Review on Deep Learning Handwritten Digit Recognition using Convolutional Neural Network. Null. https://doi.org/10.35940/ijrte.e5287.019521

Gustavo, I., Reinel Tabares-SotoRaúl, R. P., Tabares-Soto, R., Raúl, R. P., Gustavo, I., Reinel Tabares-SotoRaúl, R. P., Tabares-Soto, R., Raúl, R. P., & Gustavo, I. (2019). Deep Learning Applied to Steganalysis of Digital Images : A Systematic Review. IEEE Access, 7, 68970–68990. https://doi.org/10.1109/ACCESS.2019.2918086

Hecht-Nielsen, R. (1992). III.3 - Theory of the Backpropagation Neural Network**Based on "nonindent" by Robert Hecht-Nielsen, which appeared in Proceedings of the International Joint Conference on Neural Networks 1, 593–611, June 1989. © 1989 IEEE. (H. B. T.-N. N. for P. Wechsler (ed.); pp. 65–93). Academic Press. https://doi.org/https://doi.org/10.1016/B978-0-12-741252-8.50010-8

Hu, D., Wang, L., Jiang, W., Zheng, S., Li, B., Sangeetha, V., Prasad, K. J. R., Sedighi, V., Cogranne, R., Fridrich, J., Tabares-Soto, R., Raúl, R. P., Gustavo, I., Tan, S., Wu, W., Shao, Z., Li, Q., Li, B., Huang, J., … Chen, K. (2021). A novel steganalysis method with deep learning for different texture complexity images. IEEE Transactions on Information Forensics and Security, 16(c), 38303–38314. https://doi.org/10.1109/TIFS.2020.3025438

Hu, F., Xu, R., & Cheng, Z. (2021). A fast and effective image steganalysis model based on convolutional neural network. Journal of Physics: Conference Series, 1861(1). https://doi.org/10.1088/1742-6596/1861/1/012074

Hu, J., Hu, J., Song, W., Song, W., Zhang, W., Zhao, Y., Zhao, Y., Zhao, Y., Yilmaz, A., Yilmaz, A., Yilmaz, A., & Yilmaz, A. (2019). Deep learning for use in lumber classification tasks. Wood Science and Technology. https://doi.org/10.1007/s00226-019-01086-z

Hu, X., Chu, L., Pei, J., Liu, W., & Bian, J. (2021). Model complexity of deep learning: a survey. Knowledge and Information Systems, 63(10), 2585–2619. https://doi.org/10.1007/s10115-021-01605-0

Hussain, I., Zeng, J., Xinhong, & Tan, S. (2020). A survey on deep convolutional neural networks for image steganography and steganalysis. KSII Transactions on Internet and Information Systems, 14(3), 1228–1248. https://doi.org/10.3837/tiis.2020.03.017

Kato, H., Member, G. S., & Osuge, K. (2020). A Preprocessing by Using Multiple Steganography for Intentional Image Downsampling on CNN-Based Steganalysis. IEEE Access, 8. https://doi.org/10.1109/ACCESS.2020.3033814

Ke, Q., Ming, L. D., & Daxing, Z. (2019). Image steganalysis via multi-column convolutional neural network. International Conference on Signal Processing Proceedings, ICSP, 2018-Augus, 550–553. https://doi.org/10.1109/ICSP.2018.8652324

Li, B., Wang, M., Huang, J., & Li, X. (2014). A NEW COST FUNCTION FOR SPATIAL IMAGE STEGANOGRAPHY College of Information Engineering , Shenzhen University , Shenzhen , GD 518060 , China Institute of Computer Science and Technology , Peking University , Beijing 100871 , China. International Conference on Image Processing(ICIP), 4206–4210.

Lin, J., Ma, L., & Cui, J. (2020). A Frequency-Domain Convolutional Neural Network Architecture Based on the Frequency-Domain Randomized Offset Rectified Linear Unit and Frequency-Domain Chunk Max Pooling Method. IEEE Access, 8, 98126–98155. https://doi.org/10.1109/ACCESS.2020.2996250

Montavon, G., Samek, W., & Müller, K. R. (2018). Methods for interpreting and understanding deep neural networks. Digital Signal Processing: A Review Journal, 73, 1–15. https://doi.org/10.1016/j.dsp.2017.10.011

Okajima, Y., & Sadamasa, K. (2019). Deep neural networks constrained by decision rules. 33rd AAAI Conference on Artificial Intelligence, AAAI 2019, 31st Innovative Applications of Artificial Intelligence Conference, IAAI 2019 and the 9th AAAI Symposium on Educational Advances in Artificial Intelligence, EAAI 2019, 2496–2505. https://doi.org/10.1609/aaai.v33i01.33012496

Ranjan, R., & Forensics, C. (2016). JPEG IMAGE STEGANALYSIS USING MACHINE. IJCSI International Journal of Computer Science and Information Security, 14(2), 96–99.

Rathika, L., Loganathan, B., Scholar, M. P., Nadu, T., Nadu, T., L, R., & B, L. (2017). Approaches and Methods for Steganalysis – A Survey. Ijarcce, 6(6), 433–438. https://doi.org/10.17148/ijarcce.2017.6678

Ravi, V., & Alazab, M. (2021). Deep Learning for Cyber Security Applications : A Comprehensive Survey. ResearchGate, October. https://doi.org/10.36227/techrxiv.16748161

Sabeena, M., & Abraham, L. (2021). Digital image forensic using deep flower pollination with adaptive Harris hawk optimization. Multimedia Tools and Applications. https://doi.org/10.1007/s11042-021-10925-w

Sahu, A. K., & Sahu, M. (2020). Digital image steganography and steganalysis: A journey of the past three decades. Open Computer Science, 10(1), 296–342. https://doi.org/10.1515/comp-2020-0136

Scardapane, S., Van Vaerenbergh, S., Totaro, S., & Uncini, A. (2019). Kafnets: Kernel-based non-parametric activation functions for neural networks. Neural Networks, 110, 19–32. https://doi.org/https://doi.org/10.1016/j.neunet.2018.11.002

Shorten, C., & Khoshgoftaar, T. M. (2019). A survey on Image Data Augmentation for Deep Learning. Journal of Big Data, 6(1). https://doi.org/10.1186/s40537-019-0197-0

Tabares-Soto, R., Raúl, R. P., & Gustavo, I. (2019). Deep learning applied to steganalysis of digital images: A systematic review. IEEE Access, 7, 68970–68990. https://doi.org/10.1109/ACCESS.2019.2918086

Wang, L., Luo, Q., & Yan, S. (2022). DIESEL+: Accelerating Distributed Deep Learning Tasks on Image Datasets. IEEE Transactions on Parallel & Distributed Systems, 33(05), 1173–1184. https://doi.org/10.1109/TPDS.2021.3104252

Wang, X., Ren, H., & Wang, A. (2022). Smish: A Novel Activation Function for Deep Learning Methods. Electronics, 11(4). https://doi.org/10.3390/electronics11040540

Wang, Z., Chen, M., Yang, Y., Lei, M., & Dong, Z. (2020). Joint multi-domain feature learning for image steganalysis based on CNN. Eurasip Journal on Image and Video Processing, 2020(1). https://doi.org/10.1186/s13640-020-00513-7

Wu, Q., & Wu, Q. (2022). Research on deep learning image processing technology of second-order partial differential equations. Neural Computing and Applications. https://doi.org/10.1007/s00521-022-07017-7

Wu, S., Zhong, S., & Liu, Y. (2018). Deep residual learning for image steganalysis. Multimedia Tools and Applications, 77(9), 10437–10453. https://doi.org/10.1007/s11042-017-4440-4

Yan, X., Lu, Y., & Wang, S. (2021). Image Steganalysis via Diverse Filters and Squeeze-and-Excitation Convolutional Neural Network.

Yang, J., Shi, Y.-Q., Wong, E. K., & Kang, X. (2017). JPEG Steganalysis Based on DenseNet. 1–7. http://arxiv.org/abs/1711.09335

Yedroudj, M., Comby, F., & Chaumont, M. (2018). Yedroudj-net: An efficient cnn for spatial steganalysis. IEEE, 1, 2092–2096.

Yola, A. M., A, U. A., Gital, A. Y., Musa, K. M., & Musa, H. (2023). A New Deep Learning Network Detecting Hidden Messages in Digital Image Steganalysis. International Journal of Research Publication and Reviews, 4(11), 2892–2905.

You, W., Zhang, H., & Zhao, X. (2021). A Siamese CNN for Image Steganalysis. IEEE Transactions on Information Forensics and Security, 16, 291–306. https://doi.org/10.1109/TIFS.2020.3013204

Yu, Y., Adu, K., Tashi, N., Anokye, P., Wang, X., & Ayidzoe, M. A. (2020). RMAF: Relu-Memristor-Like Activation Function for Deep Learning. IEEE Access, 8, 72727–72741. https://api.semanticscholar.org/CorpusID:218473628

Yuan, F., Zhang, L., Wan, B., Xia, X., & Shi, J. (2019). Convolutional neural networks based on multi-scale additive merging layers for visual smoke recognition. Machine Vision and Applications, 30(2), 345–358. https://doi.org/10.1007/s00138-018-0990-3

Zeng, J., Tan, S., Liu, G., Li, B., & Huang, J. (2019). WISERNet: Wider Separate-Then-Reunion Network for Steganalysis of Color Images. IEEE Transactions on Information Forensics and Security, 14(10), 2735–2748. https://doi.org/10.1109/TIFS.2019.2904413

Zhang, C., Yu, J. J. Q., & Liu, Y. (2019). Spatial-Temporal Graph Attention Networks: A Deep Learning Approach for Traffic Forecasting. IEEE Access, 7, 166246–166256. https://doi.org/10.1109/ACCESS.2019.2953888

Zhong, S., & Chen, K. (2018). A novel steganalysis method with deep learning for different texture complexity images. Springer Science. https://doi.org/10.1007/s11042-018-6573-5 A

Zhou, X., Hong, W., Weng, S., Chen, T. S., & Chen, J. (2020). Reversible and recoverable authentication method for demosaiced images using adaptive coding technique. Journal of Information Security and Applications, 55(November), 102629. https://doi.org/10.1016/j.jisa.2020.102629

Zoph, B., & Le, Q. V. (2018). Searching for activation functions. 6th International Conference on Learning Representations, ICLR 2018 - Workshop Track Proceedings, 1–13.