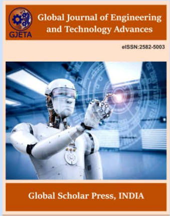
Global Journal of Engineering and Technology Advances

eISSN: 2582-5003

Cross Ref DOI: 10.30574/gjeta

Journal homepage: https://gjeta.com/




(REVIEW ARTICLE)

Check for updates

# Enhancing image authenticity: A new approach for binary fake image classification using DWT and swin transformer

Saadi Mohammed Saadi [1,*] and Waleed Al-Jawher [2]

[1] *Informatics Institute of Postgraduate Studies, Iraqi Commission for Computers and Informatics, Baghdad-Iraq.*
[2] *Uruk University, Department of Electronic and Communication, Baghdad, Iraq.*

Global Journal of Engineering and Technology Advances, 2024, 19(03), 001–010

Publication history: Received on 20 April 2024 revised on 02 June 2024; accepted on 04 June 2024

Article DOI: https://doi.org/10.30574/gjeta.2024.19.3.0091

## Abstract

The rise of social media and the easy sharing of images has led to a big increase in fake photos, making it tough to trust what we see. The proposed study, in a constructive way, presents a method of distinguishing fake from real images by the Swin Transformer and DWT (Discrete Wavelet Transform) correctly and efficiently. To capture the most features, the Swin Transformer processes the data on various levels. DWT can segment images into several frequency components required to recover different objects. It can also create a map containing all the sudden changes or discrepancies between fake images and real images by drawing features along edges and textures; It produces features across everything around the local image that are vital for detecting these things. This model is very good at tolerating noise and compression errors, making it applicable to various types of images or image manipulation. However, It focuses on local details and spatial changes. It can also pre-process and down sample image data before feeding it into Swin Transformer; possibly reducing the processing demands of the model. Using this self-attention mechanism allows the Swin Transformer to understand how different parts of an image are connected and what's going on in a bigger context. That knowledge base allows it to identify strange or unnatural patterns in its own edited photos, such as discolored areas and certain textures. Due to the hierarchical architecture of the system, it is quickly able to comprehend compounded image details and relationships which as a result has a better capability of discriminating between real and manipulated images. Putting it simply, we could come up with an attractive way to identify fake images from real ones by combining Wavelet transform and the Swin Transformer approach. This process can increase the extraction of features improve the ability and reliability of the model reduce complexity. It is an enabling technique to produce further effective solutions against the emerging problem of image manipulation and deceit in digital media. "Our research provides a strong foundation for detecting altered images and digital media, with an accuracy rate greater than 91% accuracy rate and surpassing current techniques. It can even detect minor image alterations, providing a dependable solution for digital image fraud.

**Keywords:** Classification; Image Segmentation; Discrete Wavelet Transform (DWT); Fake Images; Swin Transformer

## 1 Introduction

With a user base of over 3.8 billion people worldwide, social media platforms like Facebook, Twitter, Instagram, and Weibo have accelerated the dissemination of news and heightened public sentiment. Destructive content and false information have been disseminated by these sites[1-3]. reveals that about 64% of individuals require explanations because of misleading data. Similar research by CIGI-IPSOS and the Internet Society found that Facebook and Twitter disseminate fake news at the quickest rates[4-6]. In this paper, the Swin transformer method was used, and the reason for using it is that it works effectively as a versatile basis for computer vision, which represents a new vision. There are several differences between several fields, whether visual or verbal, whose difficulty lies in adapting a variable from

* Corresponding author: Saadi Mohammed Saadi

Copyright © 2024 Author(s) retain the copyright of this article. This article is published under the terms of the Creative Commons Attribution Liscense 4.0.

one to another. For example, the differences in the size of visual entities. When comparing images to text, we notice that images contain a much larger number of pixels compared to text, which is less [7-8]. In addition to the Swin transformer, the discrete waveform transformer (DWT) method was used in this research to divide images into frequency components, which has an effective role in making visual data patterns visible, in addition to making an analysis method to detect the frequency pattern of the image in detecting the fake image from the real one [9-10]. The main goal of the fusion method of the proposed method is accurate detection and fast processing, in addition to improving the quality of image resolution using DWT. The proposed method provides a strong framework for detecting and identifying fake images using the proposed method. The effective role was to distinguish between the changing information of the image and detect fake images from the real ones. This in turn contributes to providing safety and protecting the integrity of the visual data of the images. The proposed method has proven its methodological efficiency in enhancing efficiency and its resistance to identifying fake images, which in itself is a challenge for the field of the digital world. The research for the proposed method is organized as follows: The second section describes the methodology used to develop the classification model. Section 3 presents the results and analyses of the proposed approach, while Section 4 presents the final study observations and conclusions.

### 1.1 Related work

In this section, several studies related to this work are discussed. Gulzar, Y., in 2023, used 26,149 images of 40 different types of fruits to experiment with a modified version of MobileNetV2 architecture. The modified model, TL-MobileNetV2, achieved an accuracy of 99%, 3% higher than MobileNetV2, and a 1% error rate. Transfer learning and dropout techniques helped reduce overfitting, resulting in better results than AlexNet, VGG16, InceptionV3, and ResNet [11]. Xin Ning and others, in 2023, present an HSCF neuron model for feature processing, offering flexibility and effectiveness, and reducing dependence on training samples. It is compact hyper-sausage geometry and divisive iteration method improve classification. However, it requires iterative classifiers, larger volumes, and storage space. Future work aims to improve neurons, train on standard datasets, and explore higher-dimensional coverage geometries [12]. Juan E. Arco and others, in 2023, present a Bayesian Deep Learning-based multi-level ensemble classification system for pulmonary pathologies and Parkinson's disease diagnosis. The system maximizes performance while considering uncertainty in each decision. The system achieves an accuracy of 98.19% in pulmonary pathologies and 95.31% in Parkinson's disease diagnosis, demonstrating its applicability for clinicians due to reduced preprocessing and reliable predictions [13]. Baisen Liu and others, in 2023, introduced a new HSI classification approach using a spectral Swin Transformer network. The network enhances feature representation and contextual information, improving classification results. Experiments show accuracies of 97.46%, 99.7%, and 99.8% on public HSI datasets using the AdamW optimizer, and good generalization ability for new datasets [14]. Ruina Sun and others, in 2023, evaluated the Swin Transformer model's performance in lung cancer classification and segmentation. Pre-trained Swin-B achieved a top-1 accuracy of 82.26% in classification, outperforming ViT by 2.529%. Swin-S improved in segmentation by demonstrating mean Intersection over Union (mIoU) improvement, suggesting pre-training can enhance model accuracy [15]. T Balakrishna and others, in 2023, studied the Images that based on pixel color intensity or visual appearance, are crucial in medical imaging for managing and treating medical conditions. CT and MRI are examples of medical imaging techniques that provide comprehensive information through computer processing. The process of fusing two images using a Discrete Wavelet Transform (DWT) technique has been previously studied. This paper demonstrates the best image fusion method for a DWT technique, achieving significant detail. The method can be implemented using nine combinations of Mean-Max-Min fusion methods. The analysis and implementation of the fusion method using the DWT technique were conducted [16]. A. A. Asma and others, in 2023[17], focused on identifying Iraqi car plates to determine their location and paper origin. The new algorithm is used to identify each number on the sheet with a special color used. To analyze the image data, an evolutionary neural network was trained, and achieved good results, reaching 95 percent accuracy. The benefit of the method used is to determine the source and location of the vehicle within the framework environment In this paper, the main contribution of this study is a new idea for spotting fake images. The idea of this work is to combine discrete wavelet transform (DWT) with a Swin transformer. The combination can capture subtle variations in images to distinguish what is real from fake. It is promising, therefore, in sorting out fake images and could improve how features are extracted from images and models made adaptable and sturdy, and make models simpler. This could be the game-changer in fighting the rise of fake images and misinformation on the internet.

## 2 The Proposed Structure

The proposed framework in this study incorporates aspects of both conversion and converters to develop a usable technique for identifying fabricated images. By contrasting the terms transformation and transformer, the former alludes to the act or condition of change, whereas the latter signifies an entity that undergoes alteration and assumes a novel embodiment distinct from its preceding form. To employ the proposed structure, images are initially evaluated as genuine or manipulated (not genuine) and introduced to the system. Then, one level DWT is applied to the image to extract features from each of the four bands (LL, LH, HL, and HH). The image is then enhanced by duplicating the values of three bands (LH, HL, and HH) and keeping the original values of the LL band (approximate information of the image). The enhanced image is then fed to the binary classifier model (Swin transformer) to distinguish between real and fake images. The flow chart of the proposed Structure is shown in Figure 1, and in the following subsections, it will be explained in detail.

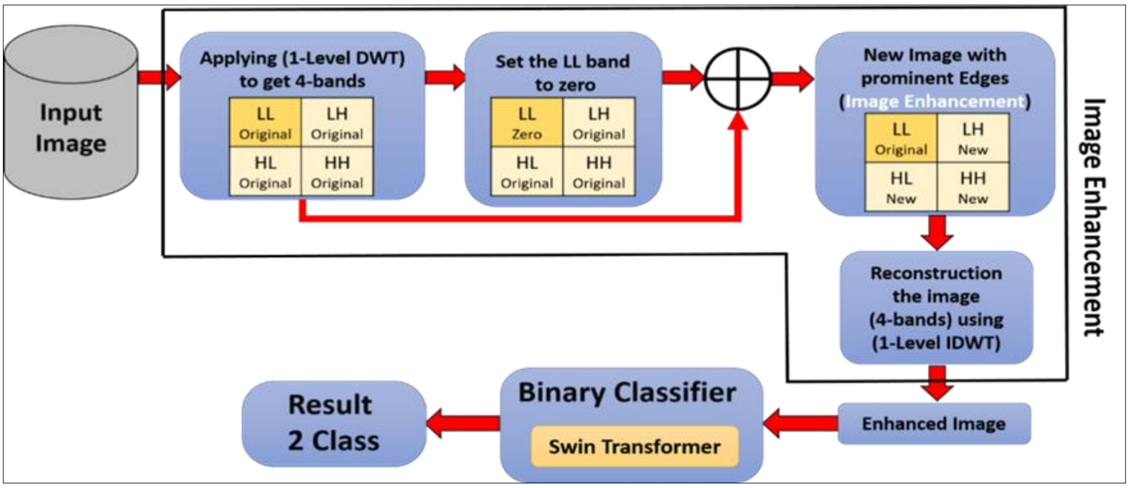



**Figure 1** Proposed Structure Flow-Chart

## 2.1 The Data Set

A dataset of images, both fake and real, categorized as either true or fraudulent, is compiled. The used dataset is on the website Kaggle, the link to the dataset is - https://www.kaggle.com/datasets/vighneshanand/oil-spill-dataset-binary-image-classification.

## 2.2 Image Enhancement

Image enhancement is the modification of digital images to make them ideal for display or further investigation. For that, one can reduce noise, enhance sharpness, or increase brightness in an image, thus facilitating the identification of important features, and there are a lot of researchers working on that using different methods and techniques in both special and frequency domains [18-19]. Direct manipulation of an image's pixels is possible with spatial domain techniques. The coordinate system of the image—also referred to as the spatial domain—is where this process takes place. However, an image is converted from the spatial domain to the frequency domain using frequency domain techniques. Mathematical transformations, including the Fourier and Wavelet transforms, are employed in this procedure. You can alter the image by adjusting its frequency components [20-21]. This study proposed, working on the frequency domain using Discrete Wavelet Transform (DWT) just like an enhancement method as shown in Figure 2:

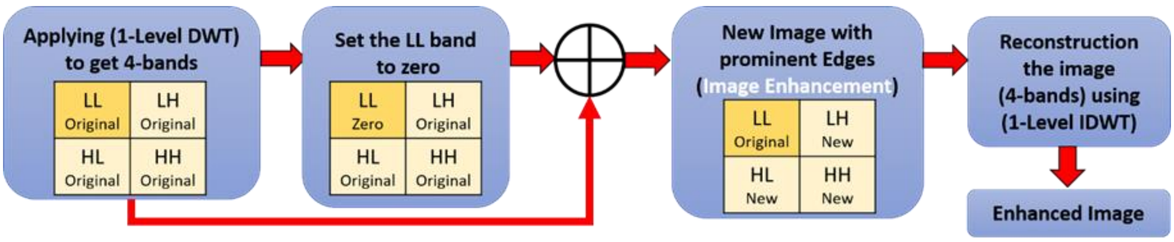



**Figure** 2 Image Enhancement

Step 1: Apply one-level Decomposition of DWT (for each channel of the image R, G, and B) and get four bands (LL, LH, HL, and HH) to extract features from an input image by using two filters (Low and High filter). The low filter passes the low frequencies (Approximate) information (the most important Feature) while the high filter passes the rest (Details) frequencies (high frequencies information).

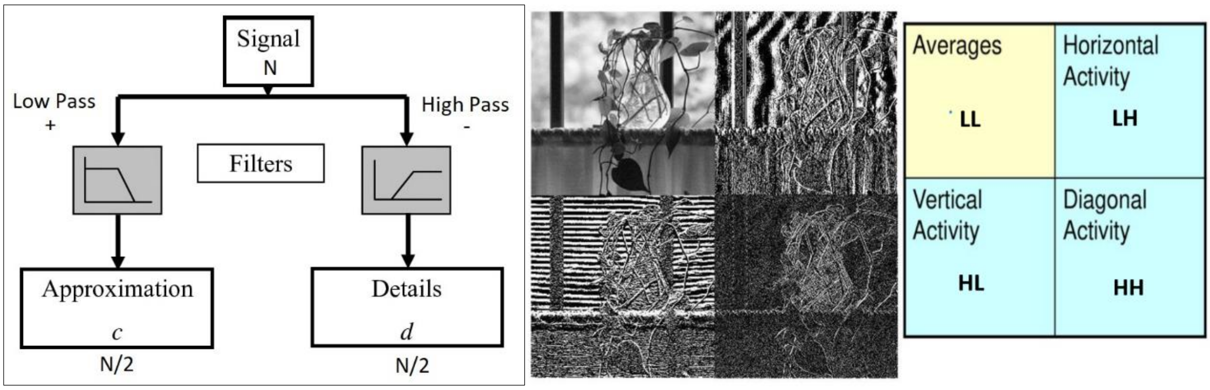



**Figure 3** Wavelet (a) Filters, (b) Four Bands

Step 2: Set the (Approximate) Low Low band (LL) to zero and duplicate the other bands (LH, HL, and HH) by two to get strong features (edges) and keep the basic and main form of the original image with enhancement.

Step 3: Reconstruct the image and return it to the spatial domain using the original values of the approximate band and the duplicated bands, to prepare it after enhancement for the classifier model.

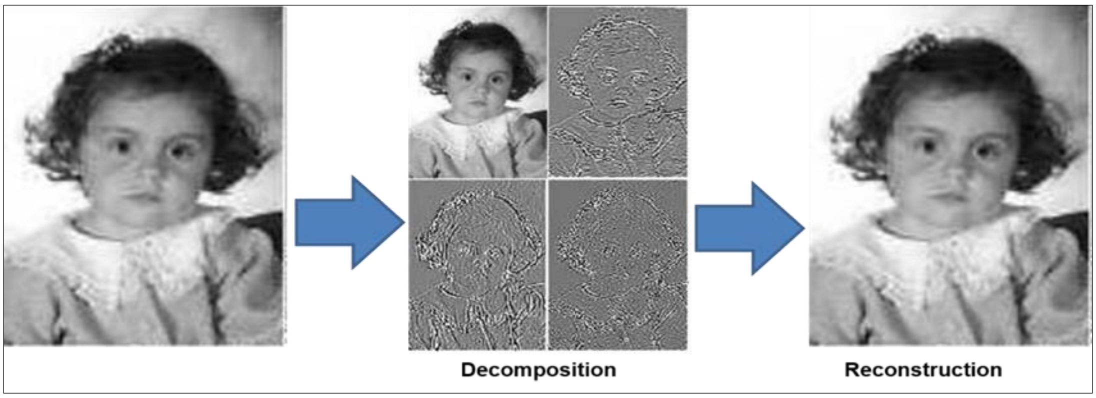



**Figure 4** Wavelet Decomposition and Reconstruction

### 2.3 The Binary Classifier using Swin Transformer

The Swin Transformer is an improved version of the Vision Transformer architecture. The model constructs hierarchical feature maps by combining picture patches in deeper layers. It achieves linear computational cost concerning the input image size since it only computes self-attention inside each local window. Therefore, it may function as a versatile foundation for both tasks involving categorizing images and ones involving accurately identifying objects. Conversely, earlier vision Transformers generate feature maps with a single, low level of detail and need quadratic computational cost concerning the input picture size, as a result of performing global self-attention computations [22-23]. The Swin Transformer architecture consists of a hierarchical structure of stages, where each stage contains a set of blocks. The blocks are composed of several layers of attention and feedforward neural networks. Figure 5 shows the architecture of the Swin transformer. The Swin Transformer consists of many components, including the Patch Partition and four stages, where in the first stage Linear Embedding and Swin transformer block are applied, and in the other three stages, Patch Merging and Swin transformer block are applied. Below is an explanation of each of these parts, the Figure 3-16 shows the applied Patch partition on the input image. [24].

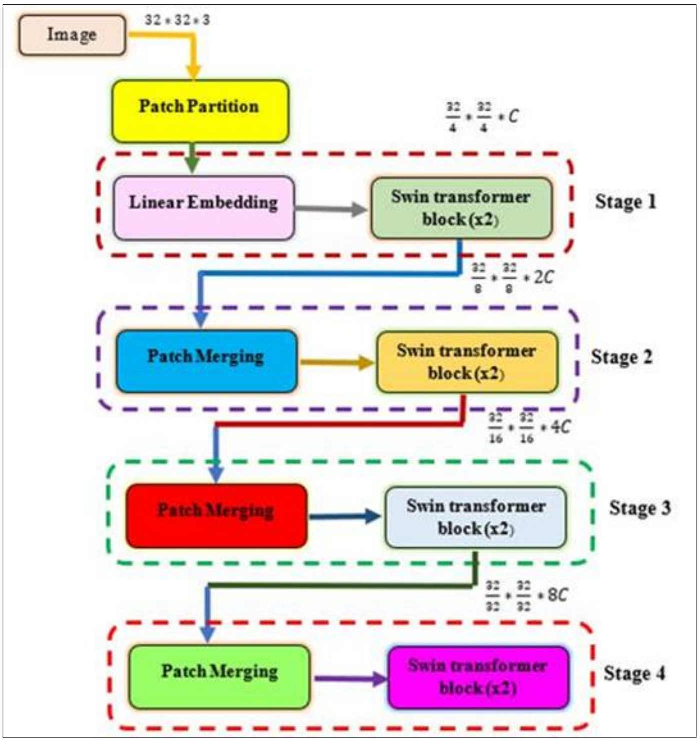



**Figure 5** The architecture of the Swin transformer

The proposed architecture in this study, resulting from the combination of DWT and Swin transformer, gave the power to detect any slight modifications in the images and thus classify them as fake and distinguish them from the original with accurate results.

## 3 Results of the proposed Structure

This section discusses the study's findings, which showed that even minute image modifications could be recognized, labeled as fake, and separated from the original with precise results using DWT [LH, HL, and HH] and Swin transformers. Precision, recall, f-measure, accuracy, sensitivity, and specificity are determined as indicated in (Eqs. 1, 2, 3, 4, 5, Eq. 6)[25–26] and will be used to evaluate a classifier.

$$Precision \ = TP/(TP + FP)(1)$$

Positive instance recall or sensitivity are highly impacted by the actual positive rate (T.P.) and false positive rate.

$$Recall \ = \frac{TP}{TP + FN} \qquad (2)$$

The accuracy, false-negative rate (F.N.), and percentage of accurate predictions are computed using the following formula.

$$Accuracy \ = (TP + FP)/(TP + TN + FP + FN)(3)$$

True negative" is denoted by "T.N.", whereas "sensitivity" refers to the quantity of positive records that yield the intended outcome.

$$Sensitivity \ = TP/TP + FN \qquad (4)$$

Accurately sorting positive records from each positive paper is what is meant by particularity.

$$Specificity \ = TN/TN + FP \qquad (5)$$

The F-measure analyzes measurements and performs numerous data recovery accuracy norms.

$$F1\ Score\ =\ 2*(Recall\ *\ Precision)\ /\ (Recall\ +Precision) \qquad (6)$$

False Negatives (F.N.) are used in erroneous classification, whilst True Positives (T.P.) and False Positives (F.P.) are used in proper classification. The sensitivity and specificity of a test define how accurately it can classify documents. The result is shown in Fig.6 Confusion Matrix for fake and real images, "fake" images are represented using class 0, and "real" images are represented using class 1. The method proposed predicted 958 images as "fake," and all of these predictions were correct (true negatives), and 1102 images as "real," and all of these predictions were correct as well (true positives). There are no false positives or false negatives, which means the model did not mistakenly label any "fake" images as "real" or any "real" images as "fake." This is an ideal result, showing 100% accuracy in this specific test dataset. It could suggest that the model is very well-tuned to the particular characteristics of the test set.

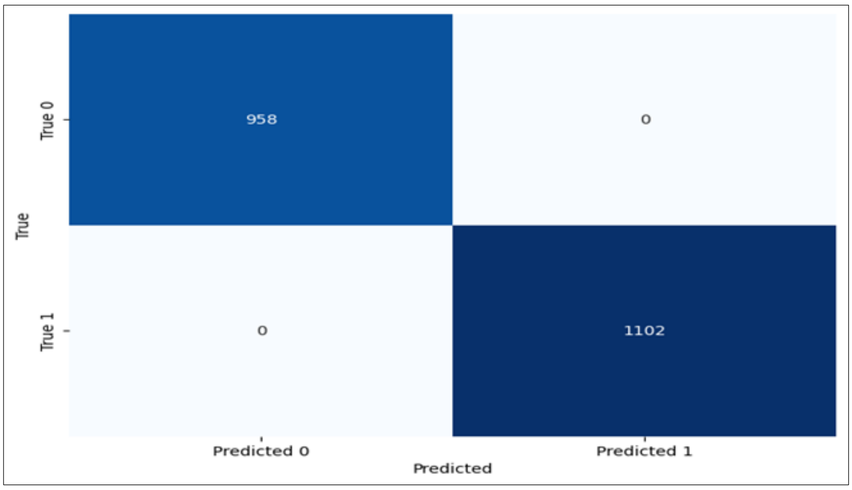



**Figure 6** Confusion Matrix for fake and real images.

The result is shown in Fig. 7. Confusion Matrix for a multi-class classification of fake and real images for 100 epochs, in which each class represents a different kind of classification for the images. The results are a summary as following:

- The cells along the diagonal (running from the top-left to the bottom-right) represent the count of accurate predictions for each class. The model has strong performance since the majority of predictions align closely with the diagonal.
- Off-diagonal cells show several misclassifications. For example, there were 2 instances where the model predicted class 1 for images that were class 0, 3 instances of class 2 predicted as class 0, and so on.

The highest misclassification seems to occur between class 2 and class 8, where the model predicted 30 instances of class 2 as class 8, and between class 8 and class 2 with 30 instances. To provide a more specific interpretation, I'd need to know what each class represents.

- The method proposed for the built model is best at predicting classes that have the highest numbers on the diagonal (e.g., class 1, class 5, class 7) and might benefit from further training or additional features to improve its predictions for classes where it's making more errors.

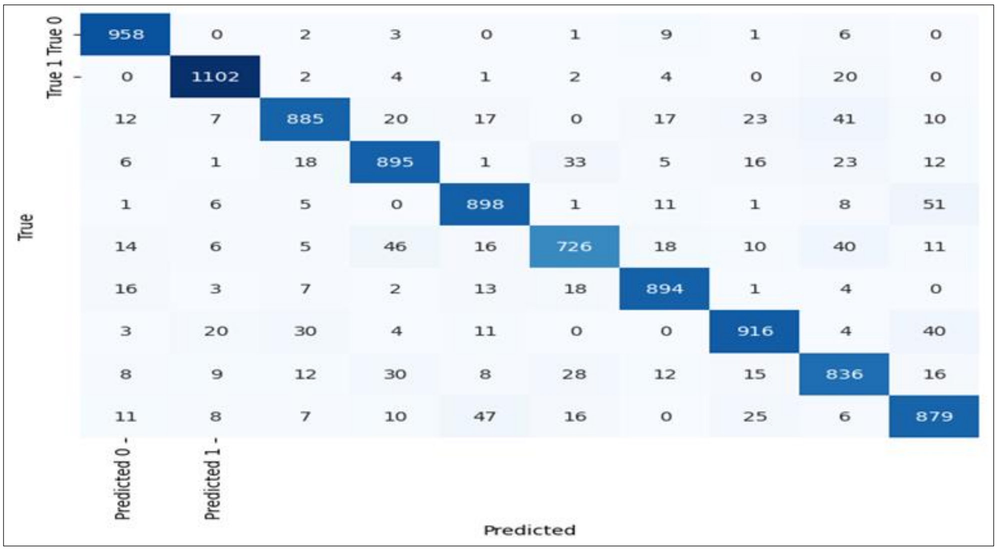



**Figure 7** Confusion Matrix for multi-class classification of fake and real images

Table 1 shows the most important results after applying the proposed method in more than one epoch. The proposed method was implemented from (10-100) and the accuracy result increased as the number of epochs increased, from 87% accuracy to 91% accuracy.

**Table 1** The result of the Proposed Model Performance of different numbers of epoch

| NO epoch | Precision | Recall | F1-score | Accuracy | Support |
|---|---|---|---|---|---|
| 10 | 87% | 87% | 87% | 87% | 10000 |
| 20 | 88% | 88% | 88% | 88% | 10000 |
| 30 | 89% | 89% | 89% | 89% | 10000 |
| 40 | 89.61% | 89.63% | 89.59% | 89.63% | 10000 |
| 50 | 89.98% | 90% | 89.96% | 90% | 10000 |
| 60 | 90.33% | 90.34% | 90.30% | 90.34% | 10000 |
| 70 | 90.43% | 90.45% | 90.42% | 90.45% | 10000 |
| 80 | 90.66% | 90.67% | 90.63% | 91% | 10000 |
| 90 | 90.77% | 90.78% | 90.75% | 90.78% | 10000 |
| 100 | 90.93% | 90.94% | 90.91% | 91% | 10000 |

A receiver operating characteristic (ROC) curve for the recommended strategy is shown in Fig. 8. As the discrimination threshold is changed, this curve shows how a binary classifier system can be diagnostic. The true positive rate (TPR), also known as recall or sensitivity, is graphed against the false positive rate (FPR) to create the curve. The FPR, which indicates specificity at various threshold values, in this graph begins at 0 and remains extremely near to 0 as the curve shifts to the right, suggesting that the classifier is not generating many false positives. This ROC curve shows the true positive rate (TPR) on the y-axis and the false positive rate (FPR) on the x-axis. When both curves in this graph reach their peaks, the TPR is 1, or 100%. The classifier represented by the dashed diagonal line, makes arbitrary estimates. usually producing an AUC of 0.5. A curve above this line should be present in any useful classifier. The class 0 and class 1 ROC curves The area under the curve (AUC) for both curves is 1.00, the highest value that can be obtained, and indicates complete discrimination between positive and negative cases for both classes. This suggests that both the true positive and false positive rates are remarkably low. The model has achieved flawless learning ability to discriminate between categories, even with highly different and separable data. If the data is overfitting, the model may not generalize well to previously unknown data since it learned the training set of data—which includes noise and outliers—too well.

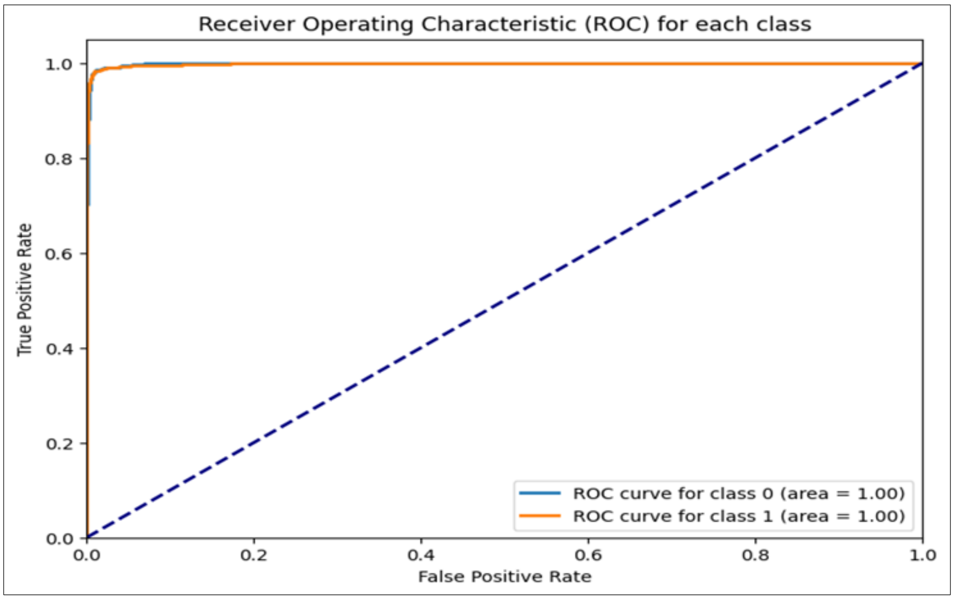



**Figure 8** A Receiver Operating Characteristic (ROC) curve for a binary classifier of fake and real images.

## 4 Conclusion

This study introduces a new Structure for identifying counterfeit images. The classification of fake and real images is achieved by applying the Discrete Wavelet Transform (DWT) and Swin transformers. To improve the image while preserving its original shape, the Low Low band (LL) can be adjusted to zero and the other bands (LH, HL, and HH) can be duplicated by two. This will result in strong features, particularly edges while maintaining the fundamental structure of the original image. After applying DWT for improved accuracy in image classification, the Swin Transformer combines the distinct shift window and global self-attention computation. The experimental results exhibit a level of performance that is comparable to other datasets. The trials revealed that the approach possessed the capability to identify even minor alterations in the photos, enabling it to accurately categorize them as counterfeit and differentiate them from the authentic ones. In the future, the suggested model will also offer a viable method for classifying deep fake models that rely on Vision Transformer.

## Compliance with ethical standards

### *Acknowledgments*

The authors would like to thank the Informatics Institute for Postgraduate Studies, Iraqi Commission for Computers and Informatics (https://iips.edu.iq/), Baghdad-Iraq, for its support in the present work.

### *Disclosure of conflict of interest*

All authors declare that they have no conflict of interest.

## References

[1] Barthel, M.; Mitchell, A.; Holcomb, J. Many Americans Believe Fake News Is Sowing Confusion. Available online: https: //www.journalism.org/2016/12/15/many-americans-believe-fake-news-is-sowing-confusion/ (accessed on 2 May 2020).

[2] Ali Adeeb, Rana, and Mahdi Mirhoseini. 2023. The Impact of Affect on the Perception of Fake News on Social Media: A Systematic Review. Social Sciences 12: 674. https://doi.org/10.3390/ socsci12120674

[3] Bago, B., Rand, D. G., & Pennycook, G. (2020). Fake news, fast and slow: Deliberation reduces belief in false (but not true) news headlines. Journal of Experimental Psychology: General, 149(8), 1608–1613. https://doi.org/10.1037/xge0000729

[4] CIGI-Ipsos Global Survey on Internet Security and Trust. 2019. Available online: https://www.cigionline.org/internet-survey-2 019 (accessed on 15 January 2021).

[5] Steinert, S. Corona and value change. The role of social media and emotional contagion. Ethics Inf. Technol. 2020, 23, 59–68. https://doi.org/10.1007/s10676-020-09545-z

[6] Stephen M. Petrie1 · T'Mir D. Julius1 , "A novel text representation which enables image classifiers to also simultaneously classify text, applied to name disambiguation", Scient metrics (2024) 129:719–743 https://doi.org/10.1007/s11192-023-04712-7.

[7] Ze Liu1,2†* Yutong Lin1,3†* and others ,Swin Transformer: Hierarchical Vision Transformer using Shifted Windows, 2021 IEEE/CVF International Conference on Computer Vision (ICCV), 2021, Pages: 9992-10002. DOI Bookmark: 10.1109/ICCV48922.2021.00986.

[8] Schmidt J, Gröller ME, Bruckner S. VAICo: visual analysis for image comparison. IEEE Trans Vis Comput Graph. 2013 Dec;19(12):2090-9. doi: 10.1109/TVCG.2013.213. PMID: 24051775.

[9] Wang Q, Li Z, Zhang S, Chi N, Dai Q. A versatile Wavelet-Enhanced CNN-Transformer for improved fluorescence microscopy image restoration. Neural Netw. 2024 Feb;170:227-241. doi: 10.1016/j.neunet.2023.11.039. Epub 2023 Nov 19. PMID: 37992510.

[10] Hao Li, Zhijing Yang, Xiaobin Hong, Ziying Zhao, Junyang Chen, Yukai Shi, Jinshan Pan, DnSwin: Toward real-world denoising via a continuous Wavelet Sliding Transformer, Knowledge-Based Systems, Volume 255, 2022, 109815, ISSN 0950-7051, https://doi.org/10.1016/j.knosys.2022.109815.

[11] Gulzar, Y. Fruit Image Classification Model Based on MobileNetV2 with Deep Transfer Learning Technique. Sustainability 2023, 15, 1906. https://doi.org/ 10.3390/su15031906.

[12] Xin Ning, Weijuan Tian, Feng He, Xiao Bai, Le Sun, Weijun Li, Hyper-sausage coverage function neuron model and learning algorithm for image classification, Pattern Recognition, Vol (136), 2023,109216, ISSN 0031-3203, https://doi.org/10.1016/j.patcog.2022.109216.

[13] Juan E. Arco, Andrés Ortiz, Javier Ramírez, Francisco J. Martínez-Murcia, Yu-Dong Zhang, Juan M. Górriz, Uncertainty-driven ensembles of multi-scale deep architectures for image classification, Information Fusion,Vol(89) ,2023, P.P.53-65,ISSN 1566-2535, https://doi.org/10.1016/j.inffus.2022.08.010.

[14] Liu, B.; Liu, Y.; Zhang,W.; Tian, Y.; Kong,W. Spectral Swin Transformer Network for Hyperspectral Image Classification. Remote Sens. 2023, 15, 3721.https://doi.org/10.3390/rs15153721.

[15] Sun, R.; Pang, Y.; Li, W. Efficient Lung Cancer Image Classification and Segmentation Algorithm Based on an Improved Swin Transformer. Electronics 2023, 12, 1024. https://doi.org/10.3390/ electronics12041024.

[16] T Balakrishna1*, Md Abdul Hannan2, G Nimishka3, and Gv Subba reddy4, DWT-Based Image Fusion Technique in Matlab: Identifying the most effective method, 4th International Conference on Design and Manufacturing Aspects for Sustainable Energy (ICMED-ICMPC 2023), Vol (391), E3S Web Conf, 2023, https://doi.org/10.1051/e3sconf/202339101076.

[17] Asma Abdulelah Abdulrahman, Fouad Shaker Tahir, Distinguishing license plate numbers using discrete wavelet transform technology-based deep learning, Indonesian Journal of Electrical Engineering and Computer Science Vol. 30, No. 3, June 2023, pp. 1771~1776 ISSN: 2502-4752, DOI: 10.11591/ijeecs.v30.i3.pp1771-1776.

[18] Marcos V. Conde, Florin Vasluianu, Javier Vazquez-Corral, Radu Timofte, Perceptual Image Enhancement for Smartphone Real-Time Applications, Computer Vision and Pattern Recognition (cs.CV); Image and Video Processing (eess. IV), Cite as: arXiv:2210.13552 [cs.CV], (or arXiv:2210.13552v2 [cs.CV] for this version),https://doi.org/10.48550/arXiv.2210.13552.

[19] Zhang, E.; Guo, L.; Guo, J.; Yan, S.; Li, X.; Kong, L. A Low-Brightness Image Enhancement Algorithm Based on Multi-Scale Fusion. Appl. Sci. 2023, 13, 10230. https://doi.org/10.3390/app131810230

[20] El Rejal, A. A., Nagaty, K., & Pester, A. (2023). An End-to-End CNN Approach for Enhancing Underwater Images Using Spatial and Frequency Domain Techniques. Acadlore Trans. Mach. Learn., 2(1), 1-12. https://doi.org/10.56578/ataiml020101.

[21] Manasha Saqib et al.,"Spatial and Frequency Domain Digital Image Watermarking Techniques for Copyright Protection", International Journal of Engineering Science and Technology (IJEST), ISSN : 0975-5462 Vol. 9 No.06 Jun 2017.

[22] Wei, C.; Ren, S.; Guo, K.; Hu, H.; Liang, J. High-Resolution Swin Transformer for Automatic Medical Image Segmentation. Sensors 2023, 23, 3420. https://doi.org/ 10.3390/s23073420.

[23] Yixuan Xu, Xianbing Wang, Hua Zhang, Hai Lin ,"SE-Swin: An improved Swin-Transfomer network of self-ensemble feature extraction framework for image retrieval", IET IMAGE PROCESSING (JAN 2024), Vol. 18, no. 1, pp. 13 – 21, https://doi.org/10.1049/ipr2.12929.

[24] Xia, H.H.; Gao, H.; Shao, H.; Gao, K.; Liu,W. Multi-Focus Microscopy Image Fusion Based on Swin Transformer Architecture. Appl. Sci. 2023, 13, 12798. https://doi.org/ 10.3390/app132312798.

[25] Saadi, S.M., Al-Jawher, W. (2024). Ensemble-based machine learning approach for detecting Arabic fake news on Twitter. Revue d'Intelligence Artificielle, Vol. 38, No. 1, pp. 25-32. https://doi.org/10.18280/ria.380103

[26] Saadi, S. M. ., & Al-Jawher, W. A. M. . (2023). Image Fake News Prediction Based on Random Forest and Gradient-boosting Methods. Journal Port Science Research, 6(4), 357–364. https://doi.org/10.36371/port.2023.4.6.