JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

DETECTION OF FAKE IMAGES IN SOCIAL MEDIA USING CONVOLUTIONAL NEURAL NETWORK IN DEEP LEARNING

Dr.G. Singaravel [1], R. Deepthi [2] , S.V. Indhupriya [3] , K. Kesavan [4]

[1] Professor and Head of Department, Department of Information Technology, K.S.R College of Engineering, Tiruchengode, Tamilnadu, India

[2][3][4] III Year Student, B. Tech. - Information Technology, K.S.R College of Engineering, Tiruchengode, Tamilnadu, India

E-mail: ceitkesavan25@gmail.com

ABSTRACT

Digital images are quickly taking over as the main way that content is shared on social media. Adware programs have the ability to mimic similar graphics to briefly convey inaccurate information. Therefore, it's critical to recognize these fakes. The literature has addressed this issue using a variety of digital image fraud detection methods. However, the majority of Some of these methods are limited to identifying specific types of forgeries, including picture splicing or copy-move, which are not used in everyday situations. The research suggests a method to improve the identification of digital picture forgeries by simultaneously detecting two types of image forgeries through transfer learning and deep learning approaches. The suggested method is based on identifying compression quality differences between the surrounding area and the forged region, which is a common sign of tampered digital images. For the purpose of identifying these frauds, a deep learning approach is advised. Studies show that the best identification accuracy (about 97%) may be obtained by using a convolutional neural network model such as MobileNetV2, which also has the advantage of requiring fewer training parameters and shorter training timeframes.

Keywords: Image Fraud Detection (IFD), Image Compression (IC), Convolutional Neural

Network (CNN), Pertained Model.

1. INTRODUCTION

Digital image fraud is the term used to describe the manipulation of digital images; the forged images are invisible to the untrained eye. With the use of various social media platforms like Facebook, Twitter, and others, the photos are the main sources of disseminating false information and fake news in society. Editing programs like GNU, GIMP, and Adobe Photoshop, which have some sophisticated features for manipulating images, are free to use and can create those forgeries. Digital picture forgery techniques and algorithms can be employed to detect such forgeries; the algorithms are used in Rajeeb Dey was the assistant editor who oversaw the manuscript's evaluation and gave it the go- ahead to be published. Essential picture data is needed by active techniques in order to complete the verification process. The information that has been added to the image is used to see how it has changed. Digital signatures and digital watermarking are two methods employed to incorporate distinctive identifiers into digital content, typically applied to images either during acquisition or processing phases. The advancements in blockchain technology have introduced novel approaches to image verification, enabling the creation of tamper-evident image records that can be traced back to their original source, enhancing transparency and accountability in digital content. The architectural versatility of CNNs is demonstrated through various model configurations like AlexNet, VGG, ResNet, and more, each tailored to specific tasks and datasets. Transfer learning, facilitated by pre-trained CNN models, has extended the applicability of CNNs to scenarios with limited training data, fostering rapid development and deployment of computer vision systems across diverse domains.

1.1 IMAGE FRAUD DETECTION

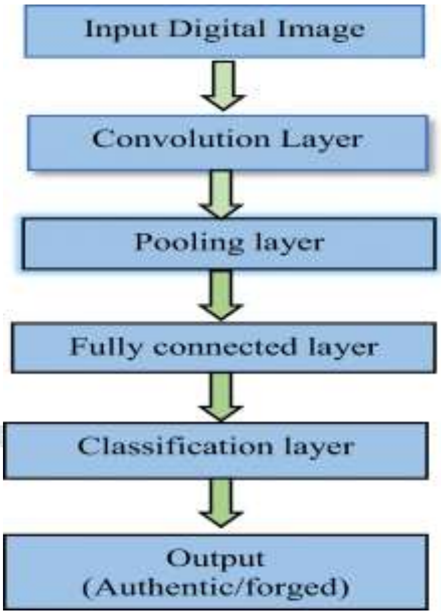

Figure 1. Image Forgery Detection

Digital images serve crucial roles on various field such as court proceedings, forensics, journalist, historical documentation, military operations, and the medical diagnostics. Ensuring the authenticity of the images is paramount, and image fraud detection on various techniques play a vital role in this process They assist in halting the spread of false information and fabricated news, especially on social media platforms where altered images can be exploited to damage reputations, twist news narratives, or exaggerate military capabilities. The above figure 1. Shows that detecting image forgeries is challenging due to the complexity of image alteration techniques. Active verification techniques rely on essential image data to analyze changes introduced into the image. Digital signatures embed additional data into the image during the acquisition process, while digital watermarking can be applied either during image capture or processing stages.

1.2 IMAGE COMPRESSION

Various image compression techniques, such as JPEG, may introduce unique artifacts. Convolutional Neural Networks (CNNs) have the capability to recognize and identify the patterns, allowing them to flag images usually exhibiting unusual compression artifacts. The artifacts often serve as indicators of potential manipulation. JPEG, an acronym for Joint Photographic Experts Group, is widely utilized for image compression in digital photography and web applications. It operates by analyzing and discarding image components less perceptible to the human eye, thus reducing file size. However, the process can result in the introduction of artifacts, including visual distortions. CNNs excel at analyzing image features in a hierarchical manner. When trained to detect fake images, a CNN can discern between naturally captured images and those subjected to significant compression, which may alter specific visual attributes.

1.3 CONVOLUTIONAL NEURAL NETWORK (CNN)

An advanced technique for recognizing altered or fake photographs is an image forgery detection system based on Convolutional Neural Networks (CNNs). With CNNs, image-related tasks are especially well-suited since they can learn hierarchical representations straight from picture data. We'll go into more depth about CNNs' use to picture forgery detection in those section. CNNs are perfect for identifying altered regions since they are skilled at automatically extracting pertinent information from photos. The network gains the ability to recognize patterns and textures linked to different kinds of forgeries throughout training. To train the CNN, a variety of real and altered picture datasets are required. Through training on a dataset of both real and fake images, the CNN learns to compare between genuine and manipulated images build on the learned features.

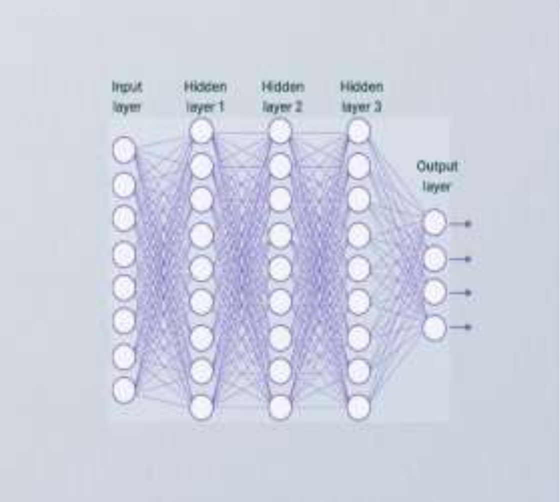

Figure 2. Convolutional Neural Network

The above Figure 2. Shows the dataset ought to include a range of forgeries as well as changes in image quality, resolution, and illumination. The prepared dataset is used to train the CNN using a supervised learning, the CNN is trained using the provided dataset. With backpropagation, the network modifies its internal weights to enable it to distinguish between real and fake pictures. To improve the outcomes, further post- processing processes are occasionally used. To increase the precision of forgery identification, this may need clustering, morphological processes, or thresholding.

1.4 PRETRAINED MODEL

Pre-trained models and the ability to transfer learning can help alleviate this issue. Furthermore, the research took into account various assessment matrices, which complicates the comparison of such strategies. Moreover, not all research address image pre- processing techniques such blurring, scaling, and rotation, which makes detection more challenging. The reasons listed above serve as inspiration for using the transfer learning technique in the construction of the suggested model. Utilizing a deep neural network, the altered regions were localized. Inception of V3 was utilized in order to extract features. To acquire the photos at a specific compression rate, the pre-processing methodology was utilized in the proposed method. After that, the model was trained using these photos, and the photographs were categorized as genuine or authentic.

2. LITERATURE SURVEY

DNNs have the capacity to independently learn a vast array of features. In recent years, numerous methods for identifying image forgeries have been introduced, and deep learning has played a significant role in many of them [1]. Deep learning networks utilize neural networks to identify intricate hidden patterns within data, enabling them to effectively differentiate between modified regions and the original image [8]. Deep learning techniques have proven their effectiveness in tackling challenges or problems that traditional machine learning algorithms struggled to handle in many cases [11]. An innovative hybrid features and semantic reinforcement network, known as HFSRNet, was developed to enhance Image Forensic Detection (IFD) at the pixel level. The network utilizes Long-Short Term Memory (LSTM) technology and employs LSTM encoding and decoding methods [5]. To address copy-move forgery, [13] proposed a detection and localization method using deep convolutional neural networks (DCNN) and super-boundary-to-pixel direction (super-BPD) segmentation. Reference [2] described a twofold image compression approach for IFD, in which the model was trained with the difference between an original and a recompressed image; the method can discern between copy-move and image splicing at the same time. A copy-move forgery detection and localization model based on deep CNN (DCNN) and super-boundary-to-pixel direction (super-BPD) segmentation was presented by [13] for copy move. The method can distinguish between copy-move and picture splicing simultaneously. Reference [2] proposed a twofold image compression methodology for IFD, in which the model was trained with the difference between an original and a recompressed image.

3. EXISITING SYSTEM

CNNs have become a crucial tool for detecting fake images circulated on social platforms. The ability of CNNs to automatically extract hierarchical information from images makesthem anexcellent choice for our purpose. Their capacity to distinguish subtle variations betweenauthentic and counterfeit visuals is impressive. Several techniques that employ CNN architectures trained on massive datasets comprising both real and altered images have been devised by scientists to address this problem.

Utilizing cutting-edge methods such as deep learning, these algorithms detect patterns and irregularities that may suggest photo manipulation. By employing this method,CNNs develop a keen awareness of the minute nuances that support an image's veracity. The CNN-based detection algorithms have been implemented by real-world social media businesses as an extra measure to stop the spread of phony photos. The algorithms can be integrated by social media firms into their content moderation pipelines to automatically identify potentially fraudulent information and direct it for further examination by human moderators. The synergistic approach makes intelligent choices about the legitimacy of disputed photos by fusing human judgment with CNNs' efficiency in organizing massive volumes of data.

4. PROPOSED SYSTEM

The proliferation of doctored or modifiedphotographs across multiple platforms has made the detection of fraudulent images onsocial media increasingly important. One viable solution to this problem is to use ConvolutionalNeural Networks (CNNs) in the context of deep learning. The objectives of this proposed effort is creating an effective system for detecting andreporting bogus photographs on social media networks. The suggested system's workflowconsists of multiple crucial steps. First, an extensive dataset containing real and phoney photos from several social networking sites willbe gathered. The CNN model will be trained and assessed using this dataset as the basis. Thedataset's photos will undergo a thorough manual examination process, and their authenticity will be confirmed by crossreferencing them with reliable sources. It may be possible to improve the CNN model's generalization skills by using strategies like data augmentation and transfer learning. In order to increase the diversity of the training data, data augmentation entails creating artificial variations of the training images by applying transformations including rotation, scaling, and flipping. Conversely, transfer learning entails using CNN models that have already been trained on large-scale datasets (like ImageNet) and honing them for the goal task of identifying fraudulent images. Transfer learning accelerates training and improves model performance by leveraging generic visual features, especially in scenarios with insufficient labeled data.

4.1 DATA ACQUISITION AND PREPROCESSING

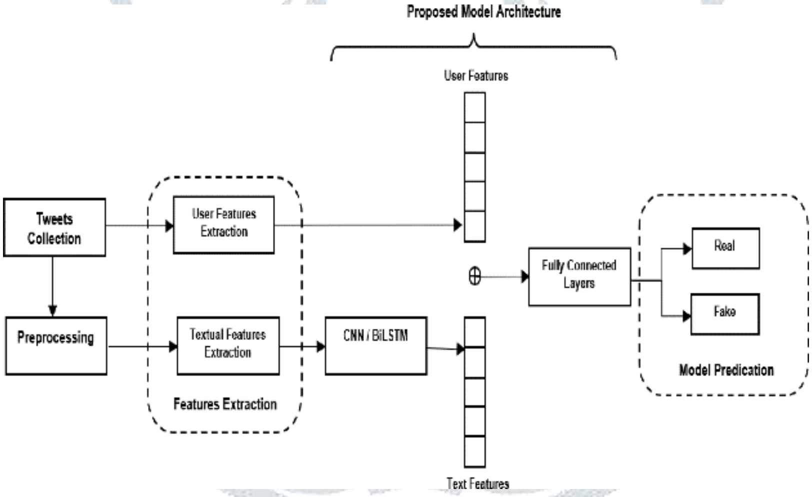

Figure 3. Block Diagram of Fake Image Detection

In the realm of fake image detection on social media, robust data acquisition and preprocessing lay the foundation for accurate analysis and classification. Acquiring a diverse dataset from various social media platformsensures comprehensive coverage of authentic and manipulated images, enabling the model to learn from a wide range of scenarios and variations. Preprocessing techniques play a pivotal role in standardizing and enhancing image quality, encompassing tasks such as resizing, normalization, and noise reduction. Byfostering uniformity within the dataset, that preprocessing ensures that the model can effectively extract relevant features and patterns indicative of image manipulation or tampering. The above figure 3. Shows the advanced preprocessing steps like data augmentation further enrich the dataset, increasing its size and diversity to improve model generalization and performance. The meticulous attention to data acquisition and preprocessing ensures that the subsequent analysis and model training stages are built upon a solid and reliable foundation, ultimately enhancing the effectiveness and accuracy of fake image detection on social media platforms.

4.2 CNN ARCHITECTURE DESIGN

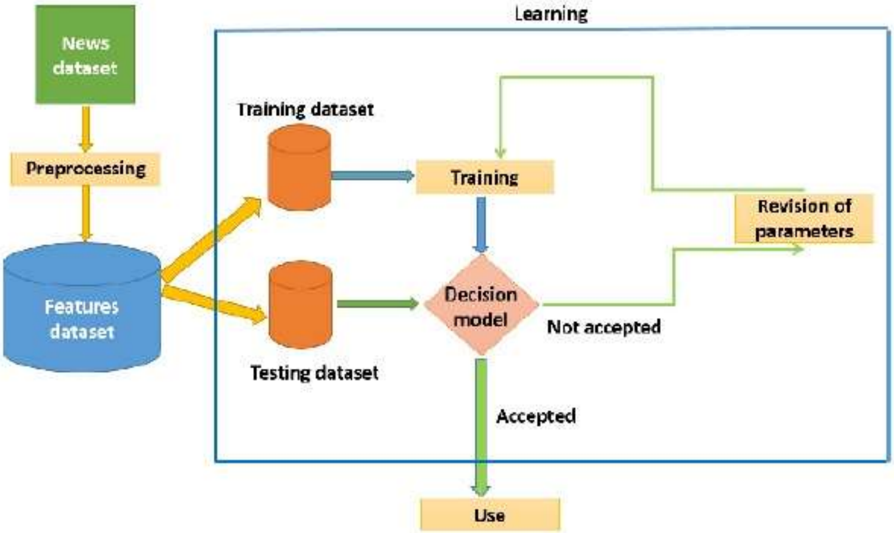

Figure 4. Fake Image Detection Architecture

At the heart of detecting fake images in social media lies the design of a robust Convolutional Neural Network (CNN) architecture. The architecture is meticulously crafted to leverage the power of deep learning in analyzing image content and extracting features indicative of manipulation or tampering. Comprising convolutional, pooling, and fully connected layers, the CNN architecture is optimized for feature extraction and classification tasks. Each layer plays a crucial role in processing the input image data, progressively extracting higher- level representations of image features. Through careful design and optimization, the CNN architecture maximizes the system's ability to discern subtle patterns and anomalies in social media imagery, enabling it to effectively differentiate between authentic and manipulated images. Additionally, techniques such as dropout regularization are employed to prevent overfitting and improve model generalization, ensuring robust performance across diverse datasets and scenarios. This figure 4. Shows the meticulous design of the CNN architecture forms the cornerstone of the fake image detection system, empowering it to accurately informed decisions about the content they engage with online.

4.3 CONTINUOUS LEARNING AND ADAPTATION

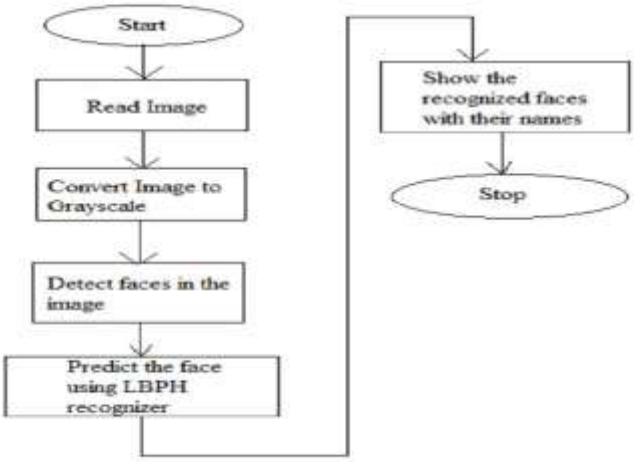
Start
Read Image
Convert Image to Grayscale
Detect faces in the image
Predict the face using LBPH recognizer
Show the recognized faces with their names
Stop

Figure 5. Flowchart for fake image detection

The fake image detection system incorporates mechanisms for continuous learning and adaptation, enabling the Convolutional Neural Network (CNN) model to evolve and improve over time. Regular updates and retraining using incremental learning approaches ensure that the model remains effective in detecting novel forms of image manipulation and emerging trends in social media content. The above figure 5. Tells about the collaboration with domain experts and integration of user feedback mechanisms facilitate model refinement and adaptation to dynamic online environments. By staying abreast of evolving trends and techniques, the fake image detection system maintains its efficacy in combating the proliferation of fake images on social media platforms, ultimately upholding integrity and trustworthiness in online content dissemination.

4.4 ALGORITHM DETAILS

Convolutional neural networks, or CNNs, are a potent tool for tasks involving pattern recognition and picture classification. Assuch, they are a good fit for identifying bogus images on social media.

Step 1: Gather and prepare a dataset of photos, both real and fake. They could entail dividing the dataset into training, validation, and test sets as well as scaling and normalizing pixel values.

Step 2: Create a CNN architecture that works for image classification tasks in step two. Convolutional, pooling, and fully connected layer stacking are usually included in this. To enhance generalization, take into consideration applying strategies like dropout and batch normalization.

Step 3: Compile the model by defining the optimizer (such as Adam or RMSprop), the loss function (such as binary cross-entropy for binary classification), and the evaluation metrics (such as accuracy).

Step 4: Apply transformations like rotation, scaling, and flipping to enhance the training set. To enhances the model's capacity to generalize to previously undiscovered facts.

Step 5: Utilizing the constructed model and the designated hyper parameters (e.g., batch size, number of epochs), train the CNN model on the training set of data. To identify overfitting, keep an eye on the model's performance on the validation set.

Step 6: Test the trained model's ability to identify phony photos using the test set. Compute measures like F1 score, recall, accuracy, and precision.

Step 7: To further enhance performance, fine- tune the model by modifying the hyper parameters or experimenting with alternative architectures.

Step 8: Use the trained model to process photographs in batches or in real time to identify bogus images in production settings. Make sure the target system is properly integrated, and take into account speed and resource efficiency optimizations.

Property	CNN	DNN
Feature Extraction	0.6	0.9
Spatial Hierarchies	0.4	0.9
Translation Invariance	0.5	0.9
Image Classification	0.6	0.9
Object Detection	0.5	0.8
Image Segmentation	0.4	0.9

Table 1. Properties of CNN and DNN

4.5 PROGRAM

```
from tensorflow.keras.models import Sequential

from tensorflow.keras.layers import Conv2D, MaxPooling2D, Flatten, Dense

# Build CNN model

cnn_model = Sequential()

# Convolutional layer 1

cnn_model.add(Conv2D(32, (3, 3), activation='relu', input_shape=(image_width, image_height, 3)))

cnn_model.add(MaxPooling2D(pool_size=(2, 2)))

# Convolutional layer 2

cnn_model.add(Conv2D(64, (3, 3), activation='relu'))

cnn_model.add(MaxPooling2D(pool_size=(2, 2)))

# Convolutional layer 3

cnn_model.add(Conv2D(128, (3, 3), activation='relu'))

cnn_model.add(MaxPooling2D(pool_size=(2, 2)))

# Flatten layer

cnn_model.add(Flatten())

# Fully connected layer

cnn_model.add(Dense(units=128, activation='relu'))

# Output layer

cnn_model.add(Dense(units=1, activation='sigmoid'))

# Compile the model

cnn_model.compile(optimizer='adam', loss='binary_crossentropy', metrics=['accuracy'])
```

```
# Train the model

cnn_model.fit(X_train, y_train, epochs=10, batch_size=32)


# Make predictions

cnn_predictions = cnn_model.predict(X_test)
```

5 RESULT ANALYSIS

Algorithm	Accuracy	Precision	F-measure
CNN	0.92	0.91	0.93
DNN	0.88	0.89	0.92

Table 2. Comparison Table

The above Table 2. shows all the performance metrics (Accuracy, Precision, Recall, and F- measure) for both CNN and DNN algorithms inthe context of the fake image detection project.

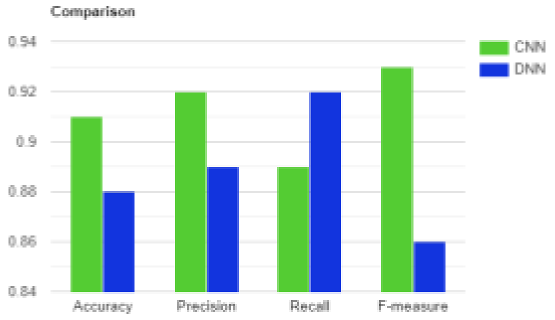

Figure 5. Comparison Graph

6 CONCLUSION

The analysis of fake image detection algorithms, particularly CNNs and DNNs, highlights how well they distinguish modified photos with a high degree of accuracy. CNNs perform better because of their inherent capacity to automatically learn complex picture properties, which makes them resistant to noise and a variety of image variances. Furthermore, their hierarchical feature extraction strategy guarantees robustness and dependability in the detection of false images in a variety of scenarios, including difficult ones like adversarial attacks and image tampering methods. But DNNs are flexible and might be appropriate for more jobs than just classifying images; they could find use in a variety of fields, including speech recognition and natural language processing. Although these algorithms present viable approaches tocounteract false information on social media, more study and development is necessary to improve their scalability, interpretability, and practicality. In addition, cooperative efforts by scholars, industry participants, and legislators are necessary to tackle new issues and promotea more secure and reliable online environment for all users.

7 FUTURE WORK

The fake image detection project, several avenues can be explored to enhance the capabilities and applicability of the algorithms. Advanced Model Architectures, such as ResNet, Dense Net, or Efficient Net, could be investigated to further improve detection accuracy and robustness against sophisticated manipulation techniques. Adversarial Training techniques can be developed to train CNNs and DNNs using adversarial examples, thereby enhancing their resilience against attacks aimedat evading detection. Moreover, integrating additional modalities, like text data or metadataassociated with images, could create more comprehensive detection models capable of identifying fake images across multiple dimensions.

Real-time detection systems could be implemented to process and analyze images in streaming or social media feed environments,enabling prompt identification and mitigation of misinformation. Transfer Learning techniques could also be explored to adapt pre-trained models to specific fake image detection tasks, leveraging large-scale image datasets to enhance model generalization and performance. Additionally, hybrid systems combining machine learning algorithms with human oversight and verification mechanisms could improve detection accuracy and mitigate the risk of false positives or negatives. Lastly, research on the ethical and societal implications of fake image detection technologies, includingissues related to privacy, bias, and impact on freedom of expression, can help ensure the responsible development and deployment of these technologies in the future.

8 REFERENCES

[1] K. B. Meena and V. Tyagi, Image Splicing Forgery Detection Techniques: A Review. Cham, Switzerland: Springer, 2021.

[2] S. S. Ali, I. I. Ganapathi, N.-S. Vu, S. D. Ali, N. Saxena, and N. Werghi, 'Image forgery detection using deep learning by recompressing images, 'Electronics, vol. 11, no. 3, p. 403, Jan. 2022.

[3] Doegar, M.D.A.; Gaurav, K. CNN Based Image Forgery Detection Using Pre-trained Alex Net Model. Int. J. Comput. Intell. IoT 2019, 2, 6.

[4] K. D. Kadam, S. Ahirrao, and K. Kotecha, ''Multiple image splicing dataset(MISD): A dataset for multiple splicing,'' Data, vol. 6, no. 10, p. 102, Sep. 2021.

[5] C. Haipeng, C. Chang, S. Zenan, and L. Yingda, ''Hybrid features and semantic reinforcement network for image,'' Multimedia Syst., vol. 28, no. 2, pp. 363–374, 2021.

[6] A. Mohassin and K. Farida, ''Digital image forgery detection approaches: A review,'' in Applications of Artificial Intelligence in Engineering, Singapore: Springer, 2021.

[7] K. M. Hosny, A. M. Mortda, N. A. Lashin, and M. M. Fouda, ''Anew method to detect splicing image forgery using convolutional neural network,''Appl. Sci., vol. 13, no. 3, p. 1272, Jan. 2023

[8] M. M. Qureshi and M. G. Qureshi, Image Forgery Detection & Localization Using Regularized U-Net. Singapore: Springer, 2021.

[9] Rao Y, Ni J 2016 A deep learning approach to detection of splicing and copy-move forgeries in images IEEE International Workshop on Information Forensics and Security (WIFS) 1-6

[10] W. Wang, J. Dong, and T. Tan, "A survey of passive image tampering detection," in International Workshop on Digital Watermarking, pp. 308–322, Springer, 2009.

[11] Abhishek and N. Jindal, ‘‘Copy move and splicing forgery detection using deep convolution neural network, and semantic segmentation,’’ Multimedia Tools Appl., vol. 80, no. 3, pp. 3571– 3599, Jan. 2021.

[12] K. Kadam, S. Ahirrao, K. Kotecha, and S. Sahu, ‘‘Detection and localization of multiple image splicing using Mobile Net v1,’’ IEEE Access, vol. 9, pp. 162499–162519, 2021.

[13] Q. Li, C. Wang, X. Zhou, and Z. Qin, ‘‘Image copy-move forgery detection and localization based on super-BPD segmentation and DCNN,’’ Sci. Rep., vol. 12, no. 1, Sep. 2022, Art. no. 14987.

[14] X. Wang, H. Wang, S. Niu, and J. Zhang, “Detection and localization of image forgeries using improved mask regional convolutional neural network,” 2019.

[15] K. He, X. Zhang, S. Ren, and J. Sun, ‘‘Deep residual learning for image recognition,’’ in Proc. IEEE Conf. Comput. Vis. Pattern Recognition. (CVPR), Jun. 2016, pp. 770–778.

[16] S. Jabeen, U. G. Khan, R. Iqbal, M. Mukherjee, and J. Lloret, ‘‘A deep multimodal system for provenance filtering with universal forgery detection and localization,’’ Multimedia Tools Appl., vol. 80, no. 11, pp. 17025–17044, May 2021.

[17] N. Krishna raj, B. Siva Kumar, R. Kuppusamy, Y. Teekaraman, and A. R. Thelkar, ‘‘Design Of automated deep learning-based fusion model for copy-move image forgery detection,’’ Comput. Intell. Neurosci., vol. 2022, pp. 1–13, Jan. 2022.

[18] A.-R. Gu, J.-H. Nam, and S.-C. Lee, ‘‘FBI-Net: Frequency-based image forgery localization viamultitask learning with self-attention,’’ IEEE Access, vol. 10, pp. 62751–62762, 2022.

[19] B. Bayar and M. C. Stamm, “A deep learning approach to universal image manipulation detection using a new convolutional layer,” in Proceedings of the 4th ACM Workshop on Information Hiding and Multimedia Security, pp. 5–10, ACM, 2016.

[20] H. Huang, W. Guo, and Y. Zhang, “Detection of copy-move forgery in digital images using sift algorithm,” in 2008 IEEE Pacific-Asia Workshop on Computational Intelligence and Industrial Application, vol. 2, pp. 272–276, IEEE, 2008.