Received: July 22, 2022. Revised: August 19, 2022.

International Journal of
Intelligent Engineering & Systems

http://www.inass.org/

IFLNET: Image Forgery Localization Using Dual Attention Network

Sunitha Krishnamurthy[1*] Krishna Alabujanahalli Neelegowda[1]
Bangalore Gnanamurthy Prasad[2]

[1]*SJB Institute of Technology, Affiliated to Visvesvaraya Technological University, Karnataka, India*
[2]*B.M.S College of Engineering, Karnataka, India*
* Corresponding author's Email: sunithakrisnamurthy@gmail.com

Abstract: The fake images and visuals can easily spread among social media users and they largely impact decision-making in society. Image forgery has become increasingly common as more non-professionals have access to image manipulation tools. These fake images are so sneaky that an ordinary person cannot guess them. Through social media, such photos are utilized to promote erroneous information in society. Image forgery detection is about segmenting the forged part from the images, primarily a region of interest. This paper suggests a unique method that depends on a dual attention network to detect forged segments. This network contains self-attention modules that contribute to extracting and matching features in the channel and spatial domains. These features help locate and identify the forged portions of digital images at various scales and channels. This experimental study uses typical datasets such as CASIA V1.0, CASIA V2.0, and Columbia. Proposed IFLNet technique outperforms other advanced techniques with a precision of 96 %, recall rate of 95 %, accuracy rate of 98 %, F1-score of 96 % and IoU score of 92 % for Columbia dataset and correspondingly other two datasets also.

Keywords: Copy-move, Splicing, Image forensics, Attention networks, Neural networks.

1. Introduction

Nowadays, altered and fraudulent pictures are shared across social media for various purposes to attract and influence people. Fake photographs are employed in various circumstances, such as journalism, police interrogation, and forensics. Editing digital photographs has grown easier due to the advent of photo editing tools. Furthermore, such software leaves no visible traces, making it extremely difficult to distinguish between a manipulated and a genuine photograph.

(a) (b)

Figure. 1 Copy move forgery technique: (a) initial image and (b) tampered image with copy move technique

(a) (b) (c)

Figure. 2 Splicing forgery techniques: (a) & b) are original images, and (c) forged images with splicing technique

Active and passive methodologies are the two fundamental sorts of tampering detection strategies. The former depends on authentication information implanted in the picture, like digital signature or digital watermark. Passively manipulated images are the most common and have been an area of examination interest. Copy-move and splicing are widely used image-altering techniques. Fig. 1 shows the altered picture with the copy-move strategy,

wherein a piece of a picture is reordered into one more picture to modify the data it contains [1], [2]. Fig. 2 shows the altered picture with splicing strategy [3], [4]; it includes replacing a particular piece of one picture with a part of another picture. Several post-processing operations [5] can also be made to the pasted region to ensure that the area created fits into the environment and is not conspicuous. These modifications include rotation, resizing, blending, and other similar techniques.

The goal of this work is to effectively localize the tampered area using convolutional neural networks (CNN). The proposed methodology uses a fully convolutional network (FCN) along with self-attention technique to describe feature interdependencies in the channel and spatial domains. DANet is a semantic scene segmentation network proposed by J. Fu [6]. This network contains two self-attention sections; namely channel attention module (CAM) and position attention module (PAM). In the spatial domain, PAM captures the spatial interdependencies across feature maps. Aggregating features use a weighted summation to update the feature at a particular point, with the weights determined by attribute matches among the matching two states. Any two states with identical attributes can help each other improve, regardless of their spatial distance. CAM is also used to capture interdependencies total station maps, with each station map receiving the burdened total of all interdependent station maps. Features obtained from both the attention modules are combined in a parallel fashion to get improved feature representations using which the forged part are localized.

This research paper is organized as follows: Section 2 illustrates the existing methods for image forgery detection and section 3 represent the suggested methodology to detect image forgery. section 4 covers detail about datasets used, and evaluation metrics and also discusses the results obtained followed by conclusions in section 5.

2. Related works

Copy-move and splicing have been the mass widely used forgery methods. Initial research works on image forgery detection used blockade methods to spot copy-move forgeries [7-13] which share the image into multiple parts. Features like discrete cosine transform (DCT), discrete wavelet transform (DWT), and local binary pattern (LBP), etc. were removed from the figure blocks and a comparison is conducted to identify similar blocks. Forgery detection is not accurate when the image's duplicated patch is post-processed and it also requires huge computational time for block comparison. Later, researchers used keypoint extraction techniques like scale invariant feature transform (SIFT), speeded up robust features (SURF), oriented FAST and rotated BRIEF (ORB) etc. to identify the forged part in an image. Keypoint-based strategy for identifying image copy-move forgeries is built using the Helmert transform and simple linear iterative clustering (SLIC) superpixel segmentation in the article [14]. The Helmert transform can also be used to find geometric correlations between matched pairs and merge clusters. SLIC approach was employed to determine the precise area of the tampering. Kunj Bihari Meena and Vipin Tyagi [15] developed a hybrid strategy by combining fourier mellin transform (FMT) and SIFT techniques. The SIFT descriptor was applied to the smooth area of the image to extract the key points from the texture area. The retrieved features are compared to find duplicate regions of the image. Under various geometric changes and post-processing operations, this approach works better in a reasonable amount of time. To further enhance the performance multiple keypoint extraction methods are used together in identifying the forged part. Keypoint based methods have the advantage of reduced computational time in forgery localization but have limitations in identifying the small and smooth forged area [16-24].

In recent times, a lot of researchers have successfully examined the usability and effectiveness of convolution neural network (CNN) [25] in the domain of image tampering, as linked to the conventional techniques. Articles [26-27], refer to a collection of significant studies in the field of image tampering detection using CNNs. Ying [28] suggest a two-stage deep learning system to detect tampering in images. A stack auto encoder network was formed based on wavelet characteristics of images to obtain intricate characteristics for each image patch. Later, the relevant data from each patch was combined, and prediction has done. A reliable identification technique must be used to deal with low-resolution images caused by compression or scaling. By detecting changes in chroma and saturation, a flat CNN has been successfully trained to distinguish tampered regions in low-resolution images in the article [29]. This has been accomplished by transforming the image from RGB to YCrCb type. CrCb channels are employed in CNN layers to eliminate the illumination details.

Researchers [30] demonstrated a new method to detect tampering using ImageNet architecture and then significantly alter the net structure using minimal training samples. The researchers suggested utilizing a coarse to refined CNN approach to detect

the modified area. A coarse CNN was utilized to forecast suspicious coarse tampering locations, and a refined CNN was used to improve the coarse-CNN identification findings [31] along with adaptive clustering technique. However, this clustering result only applies to single tampered object and may only be used to approximately localize the tamper's precise location. Researchers [32] add spatial characteristics to a U-Net encoder with a DenseNet structure to predict the binary mask of tampered areas. In uncompressed photos, the detection of tampering was lowered when using advanced tampering techniques, whereas localization of tampered areas is decreased in low-compressed JPEG images. The authors of the article [33] present an improved mask regional convolutional neural network (Mask R-CNN) that adds a Sobel filter to the Mask R-CNN to capture distinguishing features between manipulated and non-manipulated areas. The Sobel filter is used as a tool to ensure that predicted masks have image gradients that are close to those of the ground truth mask.

BusterNet technique [34] proposed a solution for detecting fake copy-move photos, which discovers the source and target areas. It also describes how well a large quantity of usable and trustworthy copy-move forgery data may be generated from datasets to address the training data deficit. Detection performance on CASIA and CoMoFoD datasets was remarkable using this approach and resilient to tampering approaches. A basic (10-layer) CNN-based deep learning technique was utilized in the article to train a hierarchical representation based on automatic RGB image creation as input. The pre-trained CNN extracts solid characteristics from the test figure, and the concluding discriminants for SVM categorization are to get used a feature fusion technique. Article [35] proposed a CNN-based approach where YCbCr, PRNU and edge features are extracted using adaptively selected ratios to produce the best-mixed features for effective tampering detection.

Trans Forensics [36] uses dense correction modules, which can correct mask predictions. This method is not only capable of capturing discriminatory plots and producing high-quality mask predictions, but it is also unaffected by patch sequence order or manipulation sorts. The encoder/decoder based technique called Fals-Unet is suggested in article [37]. Encoder, like Resnet50, analyses differences in attributes between modified and unmanipulated areas using spatial maps. To detect distorted regions, the decoder learns how to convert feature maps of low resolution into pixel-by-pixel predictions. DCU- net model includes the encoder, decoder and feature fusion [38] and can only detect splicing operations of fixed size images. Edge details of the tampered regions were retrieved using high pass filters. The tampered image, as well as the manipulated residual image, is fed into the model. The depth features retrieved from the two-channel coding network were then fused, and the modified features are extracted with varying granularities by dilation convolution, followed by secondary fusion. Finally, the decoder receives the fused feature map, to provide the expected image.

Rao [39] developed a multi-semantic attention model and integrated it into a CNN. The tamper detection performance of this approach needs to be improved for tampering in small areas. Based on the fact that JPEG files undergo double compression, Bianchi [42] proposed a method whose performance is invalid if image post processing operations, such as resizing, are used in-between the two compressions, and it occasionally generates false alarms in image regions with either low intensity. Dirik [43] and Ferrera [48] calculated the estimate of color filter array (CFA) number patterns and they analyzed noise based on CFA. FCN model [44] transformed a CNN classification into a fully convolutional classification by replacing fully connected layers in order to produce spatial heat maps. Finally, they used a deconvolution layer to up-sample the heat maps for generating dense per-pixel labeling. U-Net model [45] employed intermediate skip connection which captures low level semantic information which can be used for tampering localization.

RRU-Net [46] adds residual learning and feedback process to the traditional U-Net algorithm, which improves the detection effect of the model greatly. Lin [47] used DQ effects in the JPEG images to detect tampering without any post-processing operations. Salloum [49] evaluated a multi-task FCN (MFCN) that utilizes two output branches for multi-task learning. One branch is used to learn the surface label, while the other branch is used to learn the edge or boundary of the spliced region. Though numerous research works are done to detect and effectively localize the tampered areas, there are a few challenges that need to be addressed. This work aims to provide a framework for detecting and localizing tampered areas from digital images.

Major contributions of this work are

- Successfully capture the feature maps at different scales and channels to achieve forgery localization
- Copy-move and splicing tampering operation are detected efficiently from tampered images even

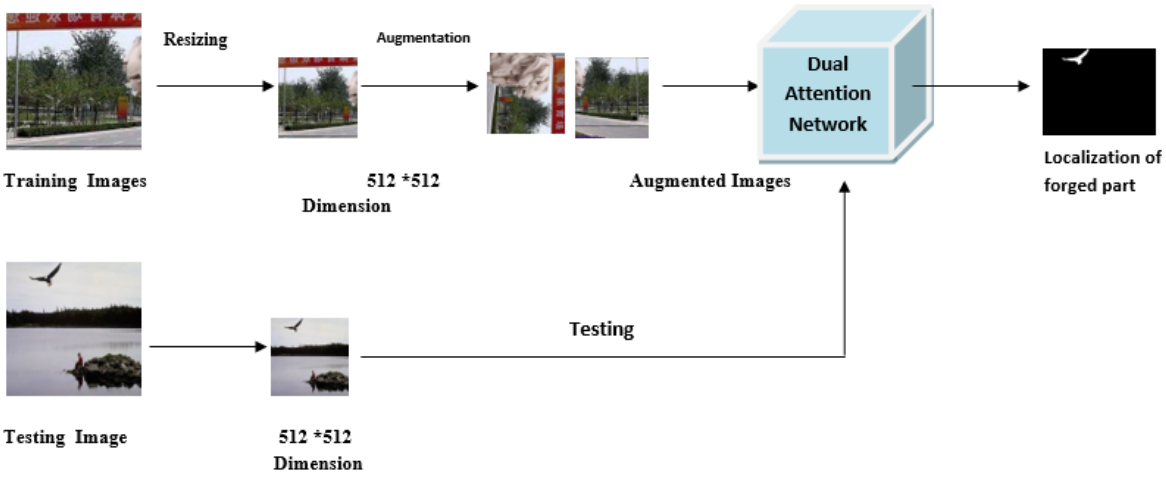

Figure. 3 Proposed methodology

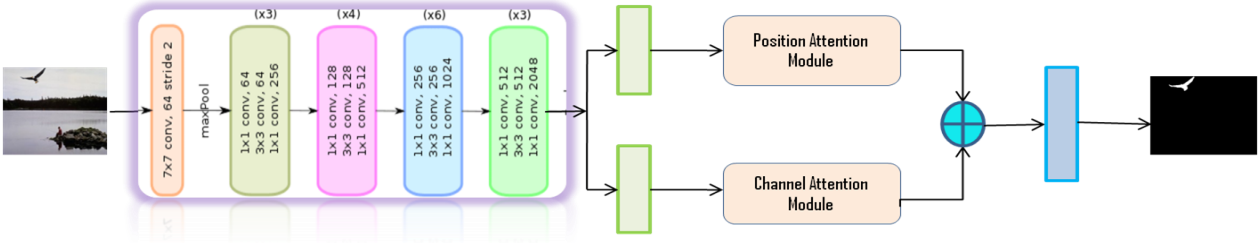

Figure. 4 DANet for image forgery detection and localization

after undergoing multiple post-processing operations.

- Achieved remarkable results in detecting tampering across data files such as CASIA V1.0, Columbia and CASIA V2.0

3. Methodology

This section discusses the working of DANet as proposed by the authors of [6], its modules, and the framework used in identifying image forgery. As shown in Fig. 3, all images from the datasets are resized to the dimension of 512x512 pixels. Images are divided in the ratio of 80:20 as training and testing images. The datasets employed in this work have fewer images, which would make the neural network's training phase less efficient. To rise above this limitation, horizontal flipping, vertical flipping, and rotation augmentation techniques are used in this work to increase the number of training and validation images. Augmented images are then passed to the DANet which extracts similar feature maps at both the channel domain and spatial domain. The ground truth image obtained from the model is in RGB format. Using masking operation, the RGB image is converted to the binary image where a region in white color indicates the forged region and other parts of the image are authentic areas. These binary images are later compared with the ground truth images for performance evaluation.

3.1 DANet model

ResNet50 is used as the backbone architecture. The last two steps of down-sampling operations are abolished, and dilated convolutions are employed, resulting in a 1/8 increase in the final feature map size. It retains a larger amount of information when no additional parameters are provided. As a result of the convolution methods applied, the features concerning pixels with similar labels will differ. Intra-class inconsistency is introduced due to these discrepancies, resulting in a loss in recognition accuracy. A connection between features and the attention mechanism is established to overcome this challenge. This technology's ability to include long-range contextual information into feature representation is useful in detecting forgeries. When combined with two different types of attention modules as depicted in Fig. 4, the network better reflects local characteristics at the pixel level while simultaneously giving a global context for the information being processed. Parallel modules would receive features from dilated ResNet. An attention matrix is created in the Position Attention module that depicts the connection among any two picture elements of the aspect. Later, it is multiplied with the

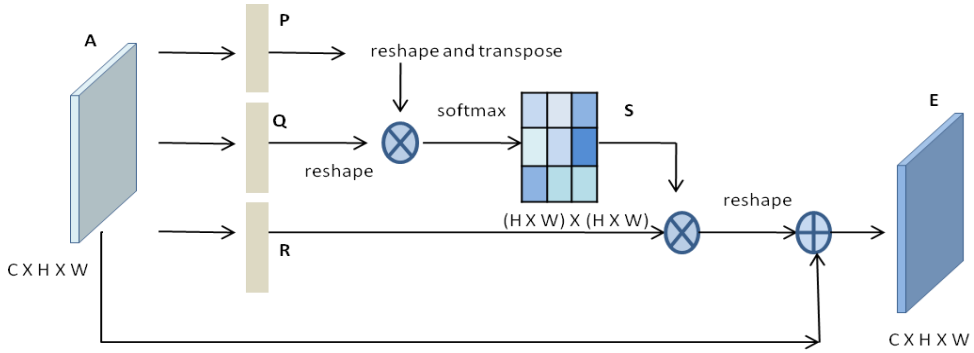

Figure. 5 PAM Network [6]

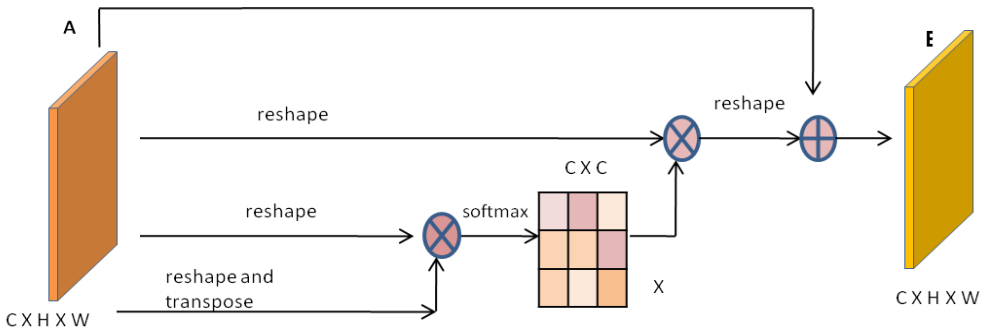

Figure. 6 CAM Network [6]

original characteristics to get the result. Long-range context representations are obtained by performing a section-wise addition process on a previously increased matrix and the original structure, described in more detail in the following section. CAM has similar steps, except for dimensionality reduction to model cross-channel relationships. The production of the two attention segment is integrated to improve the performance of forgery localization.

3.2 PAM Net

The feature of PAM is given in Fig. 5. Feature map "A" is generated from the backbone network. 1x1 Convolution is applied on $A \in \Re^{CxHxW}$ to generate P, Q and R feature maps where $(P, Q) \in \Re^{CxHxW}$. P and Q are reshaped to $\Re^{CxN}$, where N=HxW represents the sum of pixels. Matrix multiplication is performed between the transfer of P and Q. Spatial attention map $S \in \Re^{NxN}$ as shown in Eq. (1) which indicates the similar feature representations between two positions and is obtained by applying a softmax layer. The more related aspect representation of the two states provides to the greater relationship between them. R is redesigned to $R \in \Re^{CxN}$ and multiplied by the rearrange of matrix S and the product matrix is redesigned to $R \in \Re^{CxHxW}$. A section-wise addition is achieved between the product matrix and original aspect A to get the final yield $E \in \Re^{CxHxW}$ as shown in Eq. (2). Scale parameter α is primed to 0 and slowly learns to allocate more weight. Similar features would be associated between two pixels regardless of their distance. "E" has a worldwide conceptual view and aggregates settings selectively. The output feature plan of the residual network is referred to as A.

$$s_{ji} = \frac{exp(p_i, q_i)}{\sum_{i=1}^{n} exp(p_i, q_i)} \quad (1)$$

$$E^j = \propto \sum_{i=1}^{N} (s_{ji} D_i) + A_j \quad (2)$$

3.3 CAM network

The structure of CAM is shown in Fig. 6. Feature map $A \in \Re^{CxHxW}$ is redesigned to size $\Re^{CxN}$ and then matrix development is performed among A and the transfer of network attention map $X \in \Re^{CxC}$ is obtained by applying a softmax layer as shown in Eq. (3). where x_{ji} measures the i^{th} network's impact on the j^{th} network. Matrix development is achieved on the transpose of X and A and the resultant matrix is reshaped to $\Re^{CxHxW}$. The resultant matrix is grown by a scale limitation β and a section-wise sum operation is achieved with A to find the final output $E \in \Re^{CxHxW}$ as shown in Eq. (4) where β gradually acquires a weight from 0. The concluding aspect map of each network is a burdened sum of all aspects of all networks and real aspects which models the long-scope semantic addictions between aspect maps. By exploiting the interdependencies between channel

maps, we could emphasize inter-reliant aspect maps and improve the aspect demonstration of a particular definition. The yield of two attention modules is converted by a convolution layer and a section-wise sum is used to complete aspect fusion.

$$x_{ji} = \frac{A_i, A_j}{\sum_{i=1}^{c} exp(A_i, A_j)} \quad (3)$$

$$E_j = \beta \sum_{i=1}^{c} (x_{ji}. A_j) + A_j \quad (4)$$

4. Results and discussion

Various experiments are carried out for the assessment of the suggested model. The performance of the projected model in identifying and locating tampering is demonstrated using the results obtained. This section covers the datasets, and evaluation techniques.

4.1 Datasets

For evaluation, standard datasets such as CASIA V1.0, CASIA V2.0, and Columbia. The CASIA V1.0 [40] data files contain 1721 JPEG picture with an extension of 384x256 pixel sizes. The original set includes 800 photos, whereas the altered set contains 921. A genuine photograph can belong to any of the nine categories: objects, sceneries, architecture, and plants, for example. The copy and paste technique in Adobe Photoshop is used to provide tampered pictures. It may also be altered before being copied onto an image by doing a few manipulations such as rotation, scaling, and other effects.

CASIA V2.0 is a compared and expanded replica of V1.0. It has a total of 12614 different photos. The authentic set has 7,491 genuine images, whereas the altered set contains 5,123 changed images. V2.0 includes TIFF and JPEG images with varying *Q* factors. Unlike in V1.0, the images in V2.0 are of varying extent, scope from 320x240 to 800x600 picture elements. Before pasting, edited picture region(s) can be treated with ascending, rotating, or extra deformation actions to generate a combined picture. Blurring might be used on the edited region or elsewhere in the manipulated image. CASIA datasets provide ground truth images that are used for performance evaluation. Columbia dataset [41] has 183 images while the spliced set contains 180 images. Spliced images are created by copying and pasting objects of interest into actual photos using Adobe Photoshop. No post-function processes are functional to the tampered regions or images. For the Columbia dataset, border covers are provided to mark the boundaries of the manipulated area. These pictures catch whole inside or outside scenes rather than just blocks of photos which in turn is challenging.

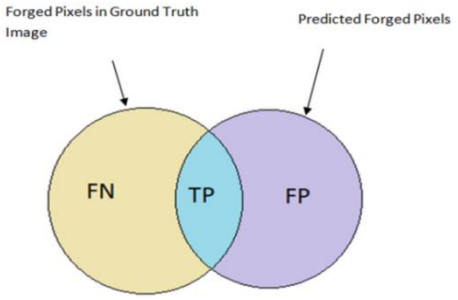

Figure. 7. Performance evaluation metrics

4.2 Evaluation metrics

Localization performance is measured at the pixel level. For evaluation, the output image generated from this work is likened to the basic facts images provided in the dataset. As illustrated in Fig. 7, T_p represents the pixels that are forged and identified as forged; F_n represents the pixels which forged but are not predicted as forged; F_p indicates non-forged pixels which are wrongly identified as forged pixels; T_n represents the non-forged pixels that are correctly categorized as non-forged pixels. Benchmark metrics such as precision P_I, intersection over union score U_I, accuracy A_I, F1-Score F_I, and recall R_I are used as shown in Eqs. (5-9). P_I refers to the chances that the identified area is truly found, while the likelihood of forged regions detected as forged is denoted by R_I. F_I represents the combined performance of P_I and R_I. Metric A_I is used to assess the accuracy of the forged area localization. The area of overlap between the predicted and actually forged segments, divided by the area of union between the two areas, is represented by U_I.

$$P_I = \frac{T_p}{(T_p + F_p)} \quad (5)$$

$$R_I = \frac{T_p}{(T_p + F_n)} \quad (6)$$

$$F_I = (2P_I R_I)/(P_I + R_I) \quad (7)$$

$$A_I = \frac{(T_p + T_n)}{(T_p + T_n + F_p + F_n)} \quad (8)$$

$$U_I = \frac{T_p}{(T_p + F_n + T_n)} \quad (9)$$

4.3 Evaluation

This section presents the outcomes of the suggested methodology and a comparison to other advanced fake detection approaches. AdamW setting is used to train the network. The localization network

Table 1. Comparative results for Columbia dataset

Method	F_I	R_I	A_I	P_I
NADQ [42]	0.2378	0.2254	0.6557	0.3292
CFA [43]	0.5836	0.5994	0.8646	0.7472
FCN [44]	0.6885	0.6126	0.8847	0.9001
C2RNet [31]	0.695	0.612	-	0.804
U-Net [45]	0.7779	0.6987	0.9134	0.985
RRU-Net [46]	0.915	0.8073	-	0.961
DU-DC-EC Net [32]	0.9307	-	0.9663	-
D-Unet [33]	0.93	0.901	-	0.96
DCU-Net-Rimg[38]	0.8858	0.8252	0.9407	0.9965
DCU-Net-RGB [38]	0.9175	0.8637	0.9545	0.9981
DCU-Net-NFF [38]	0.9216	0.9004	0.9647	0.9971
DCU-Net [38]	0.9498	0.9176	0.9727	0.9871
IFLNet (Proposed)	0.9589	0.9533	0.9212	0.96472

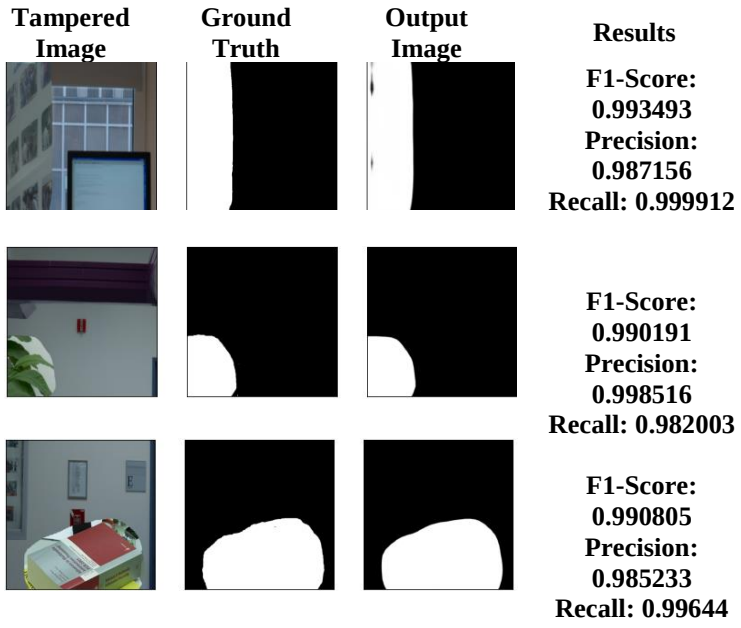

Figure. 8 Results for Columbia dataset

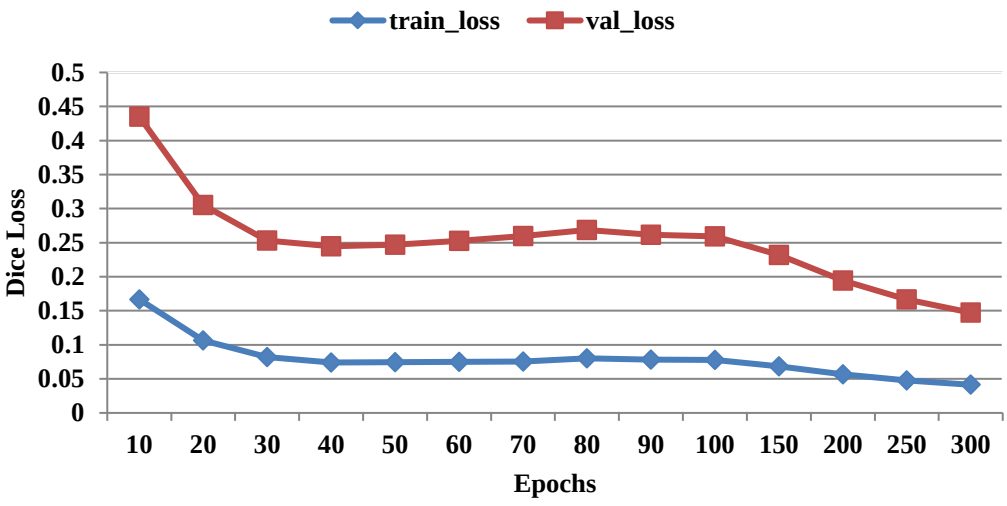

Figure. 9 Train loss vs validation loss for Columbia dataset

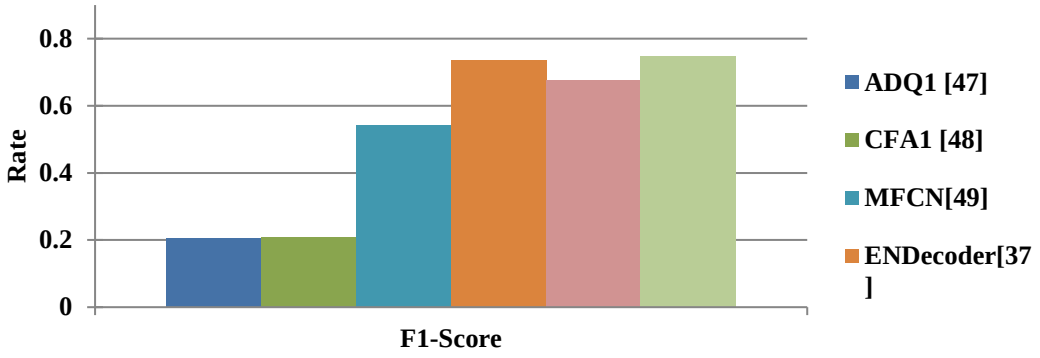

Figure. 10 Comparative analysis of CASIA V1.0 dataset

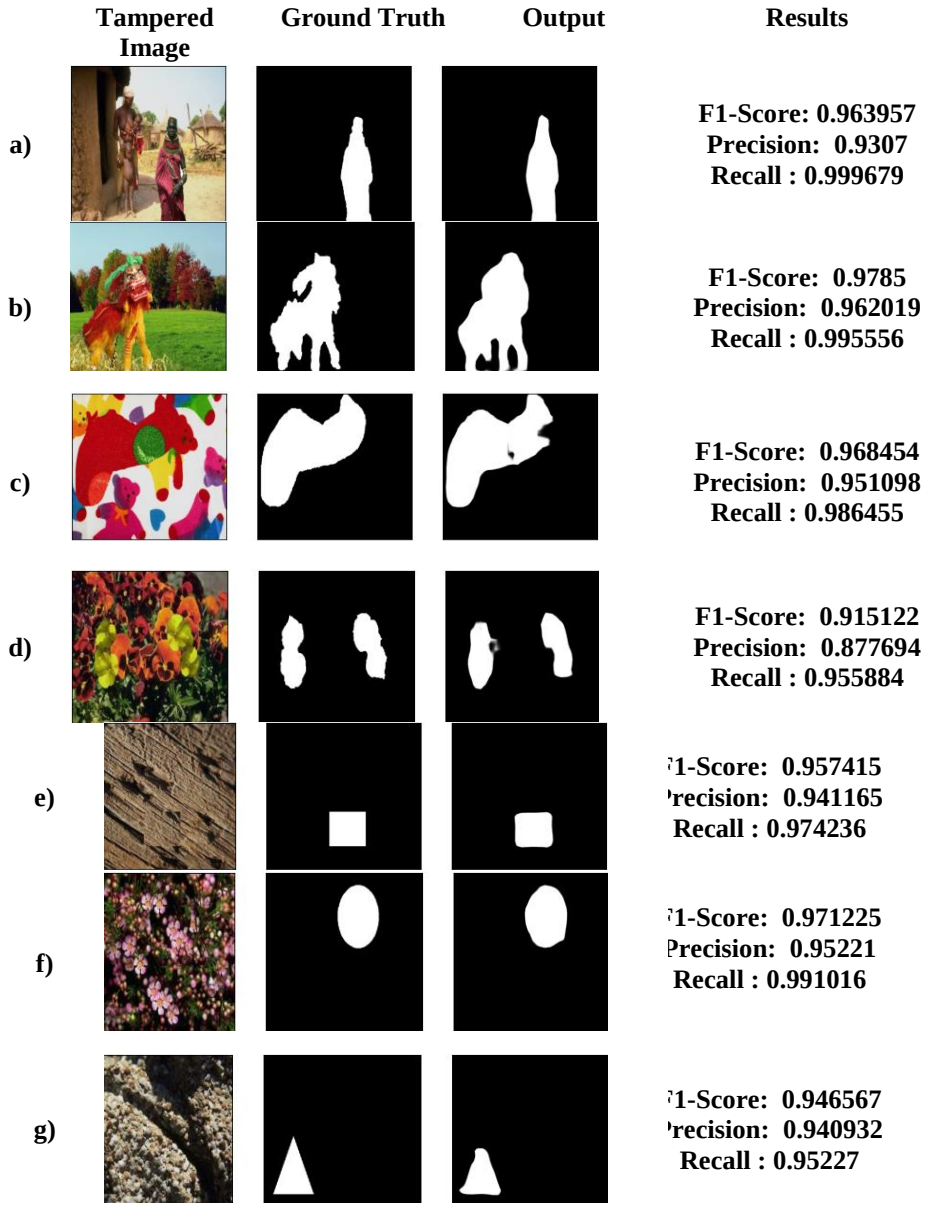

Figure. 11 CASIA V1.0 output images: (a)-(d) Detection of splicing tampered images and (e)-(g) Detection of copy-move tampered images

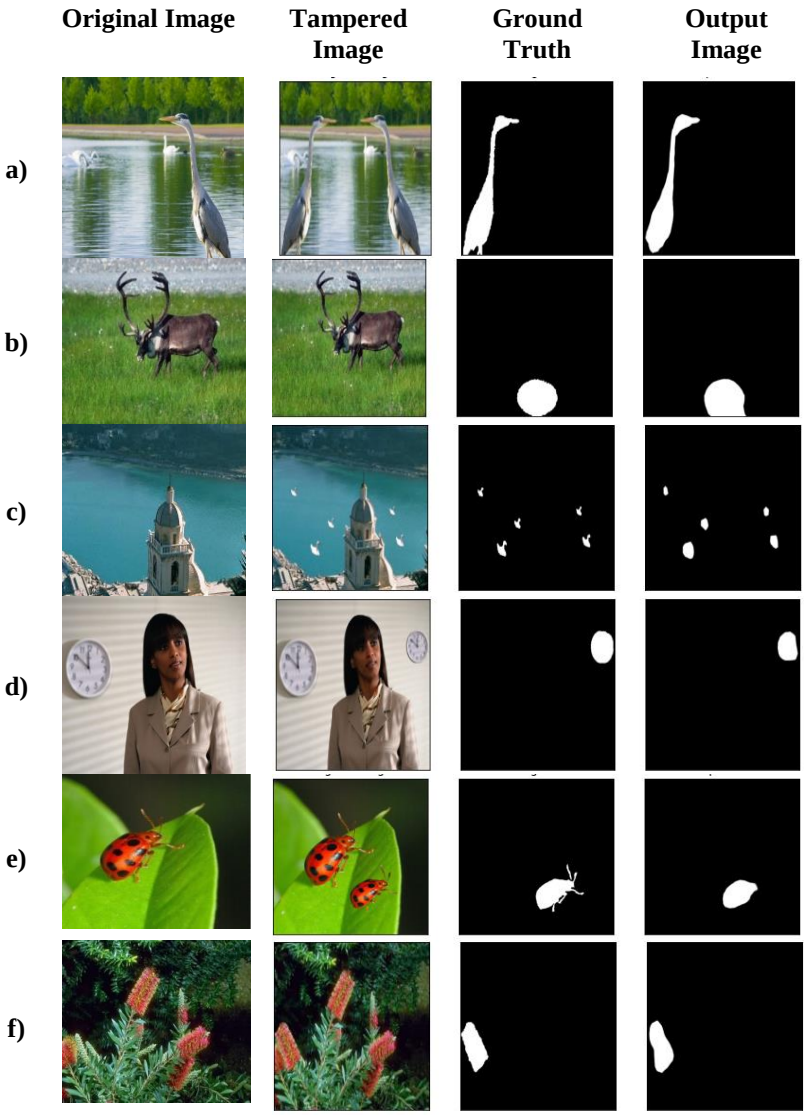

Figure. 12 Output images for CASIA V2.0 dataset

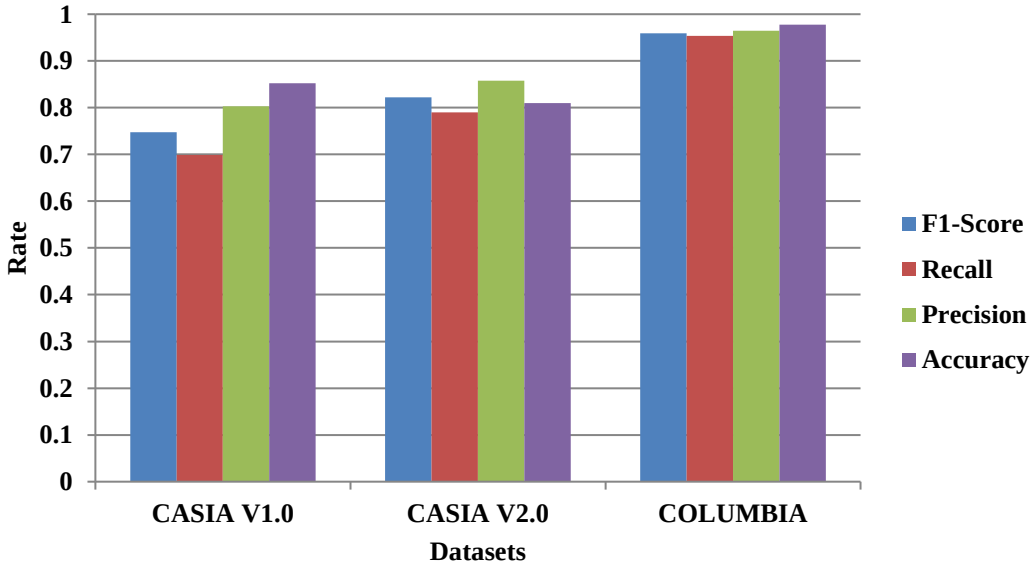

Figure. 13 Performance of CASIA V1.0, V2.0 and Columbia datasets

is defined using Pytorch 1.6.0 to train the model. Epoch value is fixed to 300, while the group size is fixed to 16. Dice loss is used for the loss purpose.

The Columbia dataset contains spliced images that are not post-processed. Fig. 8 compares the dataset's ground truth image with the predicted output image. As can be observed, the achieved results are quite comparable to the ground truth image. Fig. 9 depicts the loss during the training and validation phases. It clearly shows that the suggested model is optimal and does not undergo an over-fitting or under fitting problem. The proposed methodology's comparison study with other current approaches is shown in Table 1.

CASIA V1.0 dataset contains both spliced tampered pictures along with copy-move tampered pictures. In terms of F1-Score, Fig. 10 depicts the comparative analysis for the CASIA V1.0 dataset. Results are compared with existing work ADQ1[46], CFA1 [47], DCT [48], MFCN [49], C2RNet [31] and the result images from the suggested procedure are shown in Fig. 11. In this figure (a-d) images are spliced images and (e-g) copy-move tampered images. This figure shows that the recall rate is remarkable, implying that the forged pixels are nearly classified as forged using the proposed methodology.

CASIA V2.0 tampered photos are more difficult to work with because the tampered portions are subjected to many post-processing procedures. Even with multiple post-processing processes, the proposed methodology successfully detects tampered areas. The resulting image in Fig. 12(a) displays the tampered section copied and pasted over the same image. Detection of tampering in a textured area is shown in Fig. 12(b); the detection of multiple pasted areas and also small forged regions are shown in Fig. 12(c); the detection of tampered areas that are resized is demonstrated in Figs. 12(d), 12(e) depicts the detection of tampered regions that have undergone rotation transformation along with resizing, and similarly, Fig. 12(f) portrays the detection of the rotated tampered area.

The proposed methodology is robust to multiple transformations, copy-moves, and splicing detection in small and smooth areas. Fig. 13 depicts the overall performance of the datasets used in this work. Finding splicing operation detection from the Columbia dataset achieves better results. From the results obtained, it is understood that CASIA V1.0 is still a challenge while detecting forgeries from small and smooth regions. CASIA V2.0 dataset images contain copy move and splicing tampered areas that have undergone transformations and are highly challenging. Improving the performance of CASIA datasets would be the scope of this future work.

The proposed methodology is robust to multiple transformations, copy-moves, and splicing detection in small and smooth areas. Fig. 13 depicts the overall performance of the datasets used in this work. Finding splicing operation detection from the Columbia dataset achieves better results. From the results obtained, it is understood that CASIA V1.0 is still a challenge while detecting forgeries from small and smooth regions. CASIA V2.0 dataset images contain copy move and splicing tampered areas that have undergone transformations and are highly challenging. Improving the performance of CASIA datasets would be the scope of this future work.

5. Conclusion

The proposed method depends on the dual attention network to detect copy-move and splicing tampering detection. Resnet50 is used as a support network and two parallel attention networks (CAM and PAM) are used for feature extraction at different scale levels and channel levels. Images are augmented to provide a balanced dataset for neural network training purposes. Tampered areas are localized using this proposed methodology, and performance evaluation is done using the segmentation grade such as recall, precision, IoU score, and F1-Score. The outcomes show that the suggested procedure scores well in detecting the falsified area at a better rate than the existing other algorithms. The extensive experiments are done on Columbia, CASIA V1.0, and CASIA V2.0 datasets. The suggested methodology can detect multiple forgeries, multiple copy-move detections, and different transformations of the tampered area. Detection of regions tampered with in the case of small and smooth regions and multiple fakes detection is still a challenge for some images. Further work will consider improving the performance in detecting multiple copies and small forged areas. In the future, this research work can be used to check the performance using different attention modules to improve the feature extraction process and better segmentation of the forged area.

Conflicts of interest

The authors declare no conflict of interest.

Author contributions

The paper background work, conceptualization, methodology, dataset collection, implementation, result analysis and comparison, preparing and editing draft, visualization have been done by first author. The supervision, review of work and project

administration, have been done by second and third author.

References

[1] I. Amerini, L. Ballan, R. Caldelli, A. D. Bimbo, L. D. Tongo, and G. Serra, "Copy-move forgery detection and localization by means of robust clustering with j-linkage", *Signal Processing: Image Communication*, Vol. 28, No. 6, pp. 659–669, 2013.

[2] S. J. Ryu, M. J. Lee, and H. K. Lee, "Detection of copy-rotate-move forgery using zernike moments, in: Information Hiding", In: *Proc. of the Springer Berlin Heidelberg*, pp. 51–65, 2010.

[3] Z. He, W. Lu, W. Sun, and J. Huang, "Digital image splicing detection based on markov features in dct and dwt domain", *Pattern Recognition*, Vol. 45, No. 12, pp. 4292–4299, 2012.

[4] P. C. Man, B. Liu, X. and C. Yuan, "Multi-scale noise estimation for image splicing forgery detection", *Journal of Visual Communication and Image Representation*, Vol. 38, pp. 195-206, 2016.

[5] L. Xiang, J. H. Li, S. L. Wang, A. W. C. Liew, F. Cheng, and X. S. Huang, "Recent Advances in Passive Digital Image Security Forensics: A Brief Review", *Engineering*, Vol. 4, No. 1, pp. 29-39, 2018.

[6] Fu, J., J. Liu, H. Tian, Z. Fang, and H. Lu, "Dual Attention Network for Scene Segmentation", In: *Proc. of IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 3141-3149, 2019.

[7] Q. Wang and R. Zhang, "Double JPEG compression forensics based on a convolutional neural network", *EURASIP Journal on Information Security*, Vol. 2016, No. 1, pp. 1-2, 2016.

[8] X. Bi, C. Pun, and X. Yuan, "Multi-Level Dense Descriptor and Hierarchical Feature Matching for Copy-Move Forgery Detection", *Information Sciences*, Vol. 345, pp. 226-242, 2016.

[9] J. Zhong, Y. Gan, J. Young, L. Huang, and P. Lin, "A new block-based method for copy move forgery detection under image geometric transforms", *Multimedia Tools and Applications*, Vol. 76, No. 13, pp. 14887–14903, 2017.

[10] T. Mahmood, A. Irtaza, Z. Mehmood, and M. T. Mahmood, "Copy–move forgery detection through stationary wavelets and local binary pattern variance for forensic analysis in digital images", *Forensic Science International*, Vol. 279, pp. 8–21, 2017.

[11] J. Fridrich, D. Soukal, and J. Luka´s, "Detection of copy-move forgery in digital images", *International Journal of Computer Science*, Vol. 3, pp. 55– 61, 2003.

[12] J. Deng, J. Yang, S. Weng, G. Gu, and Z. Li, "Copy-move forgery detection robust to various transformation and degradation attacks", *KSII Transactions on Internet and Information Systems*, Vol. 12, pp. 4467–4486, 2018.

[13] D. Cozzolino, G. Poggi, and L. Verdoliva, "Efficient dense-field copy–move forgery detection", *IEEE Transactions on Information Forensics and Security*, Vol. 10, No. 11, pp. 2284–2297, 2015.

[14] H. Y. Huang and A. J. Ciou, "Copy-move forgery detection for image forensics using the superpixel segmentation and the helmert transformation", *EURASIP Journal on Image and Video Processing*, Vol. 2019, p. 68, 2019.

[15] K. B. Meenaa and V. Tyagi, "A hybrid copy-move image forgery detection technique based on fourier-mellin and scale invariant feature transforms", *Multimedia Tools and Applications*, Vol. 79, pp. 8197–8212, 2020.

[16] G. Jin and X. Wan, "An improved method for sift-based copy–move forgery detection using non-maximum value suppression and optimized j-linkage", *Signal Processing: Image Communication*, Vol. 57, pp. 113–125, 2017.

[17] M. Jaberi, G. Bebis, M. Hussain, and G. Muhammad, "Accurate and robust localization of duplicated region in copy–move image forgery", *Machine Vision and Applications*, Vol. 25, pp. 451–475, 2014.

[18] B. Shivakumar and S. Baboo, "Detection of region duplication forgery in digital images using surf", *International Journal of Computer Science Issues*, Vol. 8, No. 4, pp. 199–205, 2011.

[19] I. Amerini, L. Ballan, R. Caldelli, A. D. Bimbo, and G. Serra, "A sift-based forensic method for copy–move attack detection and transformation recovery", *IEEE Transactions on Information Forensics and Security*, Vol. 6, No. 3, pp. 1099–1110, 2011.

[20] H. Bay, T. Tuytelaars, and L. VanGool, "Surf: Speeded up robust features", In: *A. Leonardis, H. Bischof, A. Pinz (Eds.), Computer Vision – ECCV*, pp. 404–417, 2006.

[21] D. M. Uliyan, H. A. Jalab, A. W. A. Wahab, and S. Sadeghi, "Image region duplication forgery detection based on angular radial partitioning and harris key-points", *Symmetry*, Vol. 8, No. 7, p. 62, 2016.

[22] F. Yang, J. Li, W. Lu, and J. Weng, "Copy-move

forgery detection based on hybrid features", *Engineering Applications of Artificial Intelligence*, Vol. 59, pp. 73–83, 2017.

[23] K. Sunitha and A. N. Krishna, "Efficient keypoint based copy move forgery detection method using hybrid feature extraction", In: *Proc. of 2nd International Conference on Innovative Mechanisms for Industry Applications (ICIMIA)*, pp. 670–675, 2020.

[24] K. Sunitha, A. N. Krishna, and B. G. Prasad, "Copy-move tampering detection using keypoint based hybrid feature extraction and improved transformation model", *Applied Intelligence*, 2022.

[25] Y. L. Chaitra, R. Dinesh, M. T. Gopalakrishna, and B. V. Prakash, "Deep-CNNTL: Text Localization from Natural Scene Images Using Deep Convolution Neural Network with Transfer Learning", *Arabian Journal for Science and Engineering*, pp. 1-12, 2021

[26] Wu, Y., Way, A., and Rey, M. D., "Deep Matching and Validation Network An End-to-End Solution to Constrained Image Splicing Localization and Detection", In: *Proc. of the 25th ACM international conference on Multimedia*, pp. 1480-1502, 2017.

[27] B. Ahmed, T. A. Gulliver, and S. A. Zahir, "Image splicing detection using mask-rcnn", *Signal, Image and Video Processing*, Vol. 14, pp. 1035–1042, 2020.

[28] Y. Zhang, J. Goh, L. L. Win, and V. L. L. Thing, "Image region forgery detection: A deep learning approach", In: *Proc. of the Singapore Cyber-Security Conference (SG-CRC)*, pp. 1-11, 2016.

[29] Z. Zhang, Y. Zhang, Z. Zhou, and J. Luo, "Boundary-based Image Forgery Detection by Fast Shallow CNN", In: *Proc. of 24th International Conference on Pattern Recognition (ICPR)*, pp. 2658-2663, 2018.

[30] J. Ouyang, Y. Liu, and M. Liao, "Copy-move forgery detection based on deep learning", In: *Proc. of 10th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics*, pp. 1–5, 2017.

[31] B. Xiao, Y. Wei, X. Bi, W. Li, and J. Ma, "Image splicing forgery detection combining coarse to refined convolutional neural network and adaptive clustering", *Information Sciences*, Vol. 511, pp. 172–191, 2020.

[32] R. Zhang and J. Ni, "A dense u-net with cross-layer intersection for detection and localization of image forgery", In: *Proc. of IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 2982-2986 2020.

[33] X. Wang, H. Wang, S. Niu, and J. Zhang, "Detection and localization of image forgeries using improved mask regional convolutional neural network", *Mathematical Biosciences and Engineering*, Vol. 16, No. 5, pp. 4581–4593, 2019.

[34] Y. Wu, W. A. Almageed, and P. Natarajan, "Busternet: Detecting copy- move image forgery with source/target localization", *Springer International Publishing*, Vol. 2018, pp. 170–186, 2019.

[35] J. Wang, Q. Ni, G. Liu, X. Luo, and S. K. Jha, "Image splicing detection based on convolutional neural network with weight combination strategy", *Journal of Information Security and Applications*, Vol. 54, p. 102523, 2020.

[36] J. Hao, Z. Zhang, S. Yang, D. Xie, and S. Pu, "Transforensics: Image forgery localization with dense self-attention", In: *Proc. of IEEE/CVF International Conference on Computer Vision (ICCV)*, pp. 15035–15044, 2021.

[37] F. Z. E. Biach, I. Iala, H. Laanaya, and K. Minaoui, "Encoder-decoder based convolutional neural networks for image forgery detection", *Multimedia Tools and Applications*, Vol. 81, No. 16, pp. 22611-22628, 2022.

[38] H. Ding, L. Chen, Q. Tao, Z. Fu, L. Dong, and X. Cui, "Dcu-net: a dual- channel u-shaped network for image splicing forgery detection", *Neural Computing & Applications*, pp. 1–17, 2021.

[39] Y. Rao, J. Ni, and H. Xie, "Multi-semantic crf-based attention model for image forgery detection and localization", *Signal Processing*, Vol. 183, p. 108051, 2021.

[40] J. Dong, W. Wang, and T. Tan, "CASIA Image Tampering Detection Evaluation Database," In: *Proc. of IEEE China Summit and International Conference on Signal and Information Processing*, pp. 422-426, 2013.

[41] H. Y. Feng and C. S. Fu, "Detecting Image Splicing Using Geometry Invariants and Camera Characteristics Consistency", In: *Proc. of International Conference on Multimedia and Expo (ICME)*, Toronto, Canada, pp. 549-552, 2006.

[42] T. Bianchi and A. Piva, "Image forgery localization via block-grained analysis of jpeg artifacts", *IEEE Transactions on Information Forensics and Security*, Vol. 7, No. 3, pp. 1003–1017, 2012.

[43] A. E. Dirik and N. Memon, "Image tamper detection based on demosaicing artifacts", In:

Proc. of 16th IEEE International Conference on Image Processing (ICIP), pp. 1497–1500, 2009.

[44] J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation," In: *Proc. of IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Boston, pp. 3431-3440, 2015.

[45] Ronneberger, O., Fischer, P., and Brox, T. "U-Net: Convolutional Networks for Biomedical Image Segmentation", In: *Proc. of Medical Image Computing and Computer-Assisted Intervention – MICCAI, Lecture Notes in Computer Science*, Vol. 9351, 2015.

[46] X. Bi, Y. Wei, B. Xiao, and W. Li, "Rru-net: The ringed residual u-net for image splicing forgery detection", In: *Proc. of IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*, pp. 30–39, 2019.

[47] Z. Lin, J. He, X. Tang, and C. K. Tang, "Fast, automatic and fine-grained tampered jpeg image detection via dct coefficient analysis", *Pattern Recognition*, Vol. 42, No. 11, pp. 2492–2501, 2009.

[48] P. Ferrara, T. Bianchi, A. D. Rosa, and A. Piva, "Image Forgery Localization via Fine-Grained Analysis of CFA Artifacts", *In IEEE Transactions on Information Forensics and Security*, Vol. 7, No. 5, pp. 1566-1577, Oct. 2012.

[49] R. Salloum, Y. Ren, and C. C. J. Kuo, "Image splicing localization using a multi-task fully convolutional network (mfcn)", *Journal of Visual Communication and Image Representation*, Vol. 51, pp. 201–209, 2018.