Object based Image Splicing Localization using Block Artificial Grids

P N R L Chandra Sekhar[1]
Department of Computer Science and Engineering
Koneru Lakshmaiah Education Foundation
Vaddeswaram, Guntur, AP, INDIA

T N Shankar[2]
Department of Computer Science and Engineering
Koneru Lakshmaiah Education Foundation
Vaddeswaram, Guntur, AP, INDIA

***Abstract*—People share pictures freely with their loved ones and others using smartphones or social networking sites. The news industry and the court of law use the pictures as evidence for their investigation. Simultaneously, user-friendly photo editing tools alter the content of pictures and make their validity questionable. Over two decades, research work is going on in image forensics to determine the picture's trustworthiness. This paper proposes an efficient statistical method based on Block Artificial Grids in double compressed images to identify regions attacked by image manipulation. In contrast to existing approaches, the proposed approach extracts the artefacts on individual objects instead of the entire image. A localized algorithm is proposed based on the cosine dissimilarity between objects and exploit the tampered object with maximum dissimilarity among objects. The experimental results reveals that the proposed method is superior over other current methods.**

Keywords*—Image forensics; splicing localization; block artificial grids; object segmentation; double compression

I. Introduction

Now-a-days, people freely share their ideas, pictures, and comments on social networking sites. The usage of images grows enormously in different ways, such as the Government initiative towards digitizing all areas, evidence in the court of law, journalism, science, and forensics discovery [1]. Simultaneously, the widely available image editing tools induced interest in making the images or videos manipulate with ease that cannot trace out to human vision. Copy-move, splicing, resampling, cloning are few manipulation attacks to tamper images. A manipulated image significantly impacts the trustworthiness when used for evidence [2] [3]. It brings a significant challenge in image forensics to discover the original one from manipulated at the same time establish its authenticity and locate the tampered region [4].

Digital Image Forensics from Multimedia security aims at designing powerful techniques to detect manipulation attacks on images [5]. Active methods like watermarking, authentic code embedded in the original image, and verifying its authenticity. In contrast, passive methods like tampering detection do not require any external clue to assess the image's authenticity. Different tampering techniques in the literature assume that images taken from different camera models or different processing operations introduce inherent patterns into tampered image [6] [7][8] [9]. Furthermore, it assumes that these underlying patterns consistent throughout the original image, and when any manipulation attacks it, there will be inconsistency in those patterns. These inconsistency statistics can thus be used as forensic features to identify image tampering [10] [11].

In the image splicing tampering, a part of the source image is copied and pasted into the donor image. Some post-processing techniques will apply to the tampered region to make the attack invisible and difficult to trace to the human eye [12]. This challenge attracted many researchers to find various techniques for detecting image splicing. Many of these techniques extract image features and use classification to reveal for forgery, and they achieve even high success rates [13][14]. However, it is worth locating the tampered region for many real-time purposes to gain confidence. However, image splicing localization brings many more challenges as it requires pixel-level analysis rather than image-level analysis [15] [16].

The images captured by digital cameras store in the Joint Photographic Experts Group (JPEG) format. Lossy compression is used in the JPEG format and is responsible for the proliferation of images on websites and social networking sites. The image divides into 8 x 8 non-overlapping blocks in JPEG compression, and the discrete cosine transform (DCT) is evaluated for each block and then quantified using a regular quantization matrix. When any splicing attack manipulates the image, it leads to discontinuities, and these statistical traces use to exploit tampering attacks, such as JPEG quantization artefacts and JPEG grid alignment discontinuities [17] [18].

A. Related Work

The tampered blocks will undergo single compression when there is a splicing attack, while the remaining blocks will have double compression(DQ). In [19], the authors created periodic DCT patterns and evaluated each block of the image concerning its conformance of the model. Any block whose probability distribution distinguishes from the original classifies as blocks manipulated by a tampering attack. A similar approach found in [20] where the authors assume that the distribution of JPEG coefficients changes with the number of recompressions and proposes training a set of support vector machines (SVM) for the first digit artefacts and estimated the probability distribution of each block as a single or double compressed thereby exposed the splicing attack.

In [21] comparing the discontinuities using the quality factor adopted in the tampered region with the principle that a JPEG ghosts - a local spatial minimum- will correspond to the tampering attack. The limitation of the method is; it works only if the tampered region has a lower quality factor than the rest of the image. An alternative to the DQ discontinuities, in [22], the authors created a model on the entire image DCT coefficient distributions using the degree of quantization. The inconsistencies became indicative of the tampering attack. The difference between this method and the DCT-based is that the output is not probabilistic, making the technique relatively difficult to interpret although efficient.

In [17], tampering detection and localization uses the probability distribution of its DCT coefficients. Three features that can truly distinguish tampered regions from original ones are used and obtain accurate localization results. But, the refining of the probability map in post-processing influences localization results. To overcome it, [23] used a mixture model based on normalized grey level co-occurrence matrix (NGLCM) and obtained more accurate localization with the prior knowledge of both tampered and original regions. To get this, they used conditional probabilities of tampered regions and original regions of DCT blocks in first, second, and third-order statistics.

In recent works, deep-learning techniques applied for tampering detection and localized region. These methods learn the relevant features automatically from the network [24]. In [25] extracted the histograms of DCT coefficients from the input image and designed a one-dimensional convolutional neural network (CNN) with DCT coefficients as input to identify tampered regions by distinguishing single and double-compressed areas. In [26], proposed a two-layer CNN, in which the stacked auto-encoder model learns the elaborate features for the individual patch of the spliced image and uses contextual information to make the localization accurately. These methods provide block-based accuracy.

For obtaining pixel-level accuracy, [27] proposed a fully convolutional network (FCN) to locate spliced regions. FCN is a particular type of CNN, which replaces the fully connected layers with the convolutional layers having a 1x1 kernel. It distinguishes each pixel as spliced or original. The authors used three FCNs to deal with different scales of image contents, but these methods have drawbacks that they lose or smooths detailed structures and ignore small objects. To improve this effect, in [24] used a region proposal network (RPN), which is a kind of FCN and can be trained end-to-end specifically for detection. Using FCN and RPN, the authors achieved better results than FCN methods as well as other conventional methods. The computational complexity of deep-learning techniques is high.

In [28] proposes localization architecture that uses resampling features to capture artefacts. The Long short-term memory (LSTM), followed by an encoder network, is designed to differentiate tampered regions from the original. The decoder network learns features to localize the tampered region. The final soft-max layer learns the network parameters through the back-propagation algorithm from ground truth masks. The model is capable of localizing at the pixel level with high precision.

Although the deep learning-based techniques improve accuracy, they require training on large labelled databases, and the computational complexity is very high. The networks extract high-level visual features and neglect low-level features, which can be sources for forensic cues. In this paper, we move towards proposing a statistical-based forensic technique that can localize the tampered region from a single image in the presence of double compression. Unlike other techniques that produce probability maps from 8x8 DCT coefficients, we proposed an adequate statistical model that characterizes the fingerprints of block artificial grids (BAG) and works for any compression with any quality factor in the spatial domain.

B. Our Contribution

Over the years, various splicing localization techniques proposed in the literature. Still, there is scope for robustness and effectiveness to improve as splicing is complex. In this regard, we are offering the following contributions to our proposed work.

i) We propose object-based segmentation, and the features extracted from the individual objects and for each object, we estimate the variance of the BAG noise

ii) Instead of probability maps, we proposed a statistical-based localization algorithm based on pair-wise dissimilarity among objects to classify the suspicious object from the original ones.

The rest of the paper organizes as follows: Section II described JPEG fingerprints from block artificial grids to speed up computation time. Section III outlines the proposed statistical method to expose and localize the splicing attack. The experimental and evaluation results present in Section IV, and finally, the paper concluded in Section V.

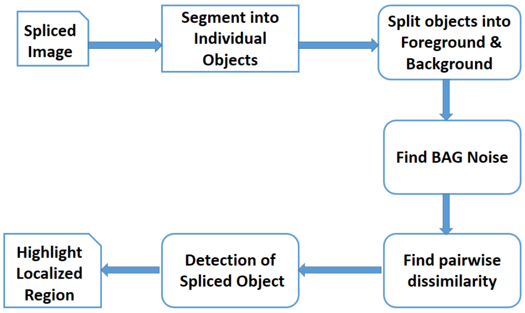
Spliced Image
Segment into Individual Objects
Split objects into Foreground & Background
Find BAG Noise
Find pairwise dissimilarity
Detection of Spliced Object
Highlight Localized Region

Fig. 1. The Proposed Frame Work

II. Proposed Method

The primary goal is to localize the tampered region in the spliced image. As shown in Fig. 1, the proposed method is in three levels: object-based image segmentation to extract individual objects from the spliced image and estimate each object's variance using block-artificial grids and the proposed localization algorithm on pair-wise dissimilarity among objects to expose tampered region.

A. Object Segmentation

Object Detection is a complicated computer vision problem to detect and classify objects from an individual image or videos. In many existing popular object detection frameworks, Mask R-CNN [29] is a frequently used one developed by Facebook research. It is an extension of Faster R-CNN that

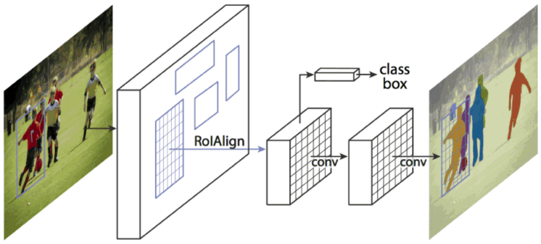

Fig. 2. Mask R-CNN Frame Work adopted from [29]

estimates the object's mask and human pose. It overcomes the COCO suite challenge by segmentation of instances, detecting bounding-box objects, and individual key points.

Using the Mask R-CNN framework, as shown in Fig. 2, performed object detection and segmentation [30] for the given spliced image extracted individual masks of all objects. Then for each mask, find its object from the input image along with the bounding box area. The object corresponds to the mask considered a foreground object, and the remaining part in the bounding box region is the background object.

B. *Block Artificial Grids*

The lossy JPEG compressed image leaves horizontal and vertical breaks in the image and is commonly refers to as Block Artificial Grids (BAG). The image's BAGs are roughly at the border of a 8 x 8 block with a periodicity of 8 at both horizontal and vertical edges. When any manipulation attack alters the image, the BAGs appear within the block instead of at borders. Thus this JPEG fingerprint is used in image forensics [31].

While compress the image using a digital camera, it introduces noise such as natural noise, BAG noise due to the JPEG compression factor. The artificial grid lines in a 8 X 8 block are feeble than the border edges. In [31], the authors extracted weak horizontal and vertical lines of a grayscale image with a periodicity of 8 separately to enhance these weak lines, and then combined them is referred to as BAGs.

In this paper, we focus on extracting BAGs in colour images. Since the luminance component in the JPEG standard is 8 x 8 blocks, we used only the luminance component rather than C_b and C_r of components of the YC_bC_r image.

The second-order difference of an image regards as weak horizontal edges of an image. For the given image $I(m,n)$, the absolute second-order difference $d(m,n)$ is obtained by

$$d(m,n) = |2I(m,n) - I(m+1,n) - I(m-1,n)| \tag{1}$$

A median filter is applied to enhance the weak edges and remove the interference coming from strong image edges. To further reduce the edge influence as in [31] ignored differentials greater than an experimental threshold. Then the enlarged horizontal edges are accumulated for every two subsequent blocks as:

$$e(m,n) = \Sigma_{i=n-16}^{16} d(m,i) \tag{2}$$

Then to equalize the amplitudes throughout the resultant image, a local median is reduced from each element.

$$e_r(m,n) = e(m,n) - median[\{e(i,n)|m-16 \leq i \leq m+16\}] \tag{3}$$

Thus, the weak horizontal edge image w_h obtained by applying the periodical median filter as:

$$w_h(m,n) = median[\{e_r(i,n)|i = m-16, m-8, m, m+8, m+16\}] \tag{4}$$

where $w_h(m,n)$ are elements of extracted horizontal BAG lines. The five elements in Eq. 4, with spacing eight used in the median filter, makes the strong BAGs and weak BAGs smooth, and rest are removed. As more elements used in the median filter, BAGs can extract in a better way.

The vertical BAGs w_v are also similarly extracted.

$$w_v(m,n) = median[\{e_r(m,i)|i = n-16, n-8, n, n+8, n+16\}] \tag{5}$$

The final BAG obtained by combining Eq. 4 and 5 as

$$w_b(m,n) = w_h(m,n) + w_v(m,n) \tag{6}$$

Eq. 6 gives BAGs for the original image. In the tampered image, the BAGs appear at some abnormal position, such as the block center. So, for a fixed 8 x 8 block w_mn, these abnormal BAGs can be obtained as [31].

$$\begin{aligned} w_mn = & Max\{\Sigma_{i=2}^{7} w_b(i,n)|2 \leq n \leq 7\} \\ & - Min\{\Sigma_{i=2}^{7} w_b(i,n)|n = 1,8\} \\ & + Max\{\Sigma_{i=2}^{7} w_b(m,i)|2 \leq m \leq 7\} \\ & - Min\{\Sigma_{i=2}^{7} w_b(m,i)|m = 1,8 \end{aligned} \tag{7}$$

C. *Localization of Splicing Region*

Mask R-CNN object detection framework [30] is used to detect individual masks from the spliced image. For each mask, first split into the foreground and background objects and extracted the BAGs, as discussed in Section II-B.

To expose discrepancies in BAGs of individual objects, we find BAG noise from Eq. 7 as:

$$\mu = \frac{1}{R}\Sigma w_mn(i,j) \sigma = \frac{1}{R}\Sigma(w_mn(i,j) - \mu)^2 \tag{8}$$

μ is mean, σ is variance, and R represents the no of BAG features in w_mn.

After BAG noise obtained for each object, pair-wise dissimilarity among objects evaluated as follows:

For each pair of the distinct foreground or back-ground objects, let the BAG noise be S_1 and S_2. Then the cosine dissimilarity between the objects defined as:

$$L_D = 1 - \frac{C(S_1, S_2) + 1.0}{2} \tag{9}$$

where

$$C(S_1, S_2) = \frac{S_1^T . S_2}{\|S_1\| . \|S_2\|} \tag{10}$$

$C(S_1, S_2)$ is the cosine angle between two BAG noises. The metric L_D gives values in the range [0,1]. Where the values near to 0 represent similar BAG noise levels of both objects, and near to 1 represents different levels.

Algorithm 1 Algorithm for identifying probable tampered object from Dissimilarity Matrix

```
Input: Estimated noise levels of $N$ Individual objects of Spliced Image
Output: Tampered Object Find Dissimilarity matrix
for $i = 1$ to $N$ do
    for $j = 1$ to $i - 1$ do
        $DM(i, j) = L_D(S_i, S_j)$
    end for
end for
Find the pair having maximum dissimilarity
for each column in $DM$ do
    $[COLMAX_j, COLIDX_j] = max(DM_j)$
end for
$[cmax, cidx] = max(COLMAX)$
for each row in $DM$ do
    $[ROWMAX_i, ROWIDX_i] = max(DM_i)$
end for
$[rmax, ridx] = max(ROWMAX)$
$DM(rmax, cmax)$ has maximum dissimilarity
Now find which object has maximum dissimilarity
$RROW = ridx, count = 0$
for each $COLIDX$ do
    if $(RROW = COLIDX_j)$ then
        $count = count + 1$
    end if
end for
 retrun the tampered object
if $(count \geq 0)$ then
    $T_P = RROW$
end if
return $T_P$
```

The probable tampered object with maximum dissimilarity with other objects is exposed from the dissimilarity matrix using the proposed localization algorithm 1.

III. Experimental and Performance Analysis

This section evaluates the proposed method on two datasets and compares its performance with contemporary techniques.

Typically, CASIA dataset [32] is a widely used evaluation dataset for JPEG image splicing forgery detection, and it consists of 7491 authentic and 5123 spliced images with JPEG, TIFF, and BMP types of images. We randomly selected 1000 tampered images of animals, persons, birds, vehicles with the size 384 x 256 and segmented the objects using the Mask R-CNN framework. The proposed method is tested on those chosen tampered images of the CASIA dataset for localizing spliced regions.

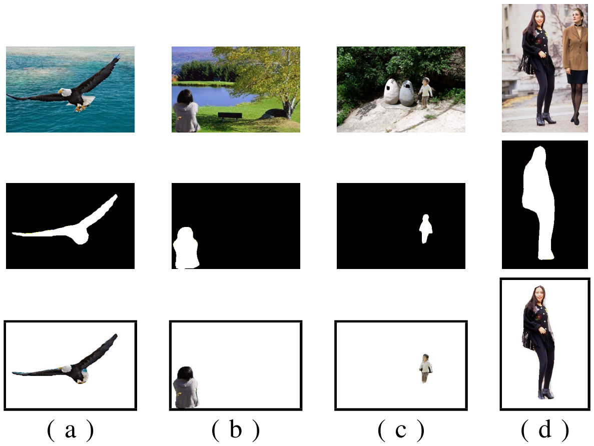

Fig. 3. Visual Evaluation of Proposed Method on CASIA Dataset

The qualitative evaluation of splicing images on the CASIA dataset shows in Fig. 3. The first row consists of randomly chosen four images, and the respective ground truth masks given in the second row. The proposed method results are in the last row, where the spliced region is highlighted, and the remaining area is marked as white. From the results, the proposed method's superiority is very clearly evident to localize the spliced region.

To increase the proposed method's robustness, we have evaluated our approach on the Image Manipulation Dataset (IMD) [33]. The dataset contains a 48 pixel high-resolution JPEG compressed images with size 3264 x 2448 with different quality factors ranging from 20% to 100%. The images were cropped to 2048 x 1536 to reduce the computational complexity and spliced each other and obtain 600 spliced images. Then the proposed method was assessed on those images.

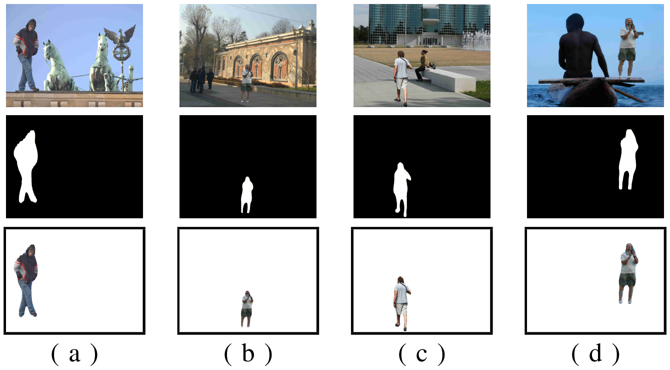

Fig. 4. Visual Evaluation of Proposed Method on the High Resolutioned Images from [33]

The evaluation results on our customized IMD spliced dataset obtained from [33], shown in Fig. 4. The first row contains randomly chosen four sample images from the dataset. The ground truth masks are in the second row, and the proposed method results are in the third row. From the results, the proposed method works well on high-resolution images.

A. Localization Accuracy

The accuracy of splicing localization evaluates based on pixel-level F-measure. Two metrics, True Positive Rate (TPR), measure the rate of pixels that are indeed detected as spliced, and False Positive Rate (FPR), a measure of the rate of pixels that are falsely detected as spliced, are used to evaluate F-measure.

$$TPR = \frac{TP}{TP + FN} * 100 FPR = \frac{FP}{FP + TN} * 100 \quad (11)$$

Where TP is True Positive, FP is False Positive, TN is True Negative, and FN is False Negative. It expects to have high TPR and low FPR in the results. From these metrics, the F-measure defines as follows:

$$F = 2 * \frac{TPR * FPR}{TPR + FPR} \quad (12)$$

We evaluated average TPR and FPR and F-measure for all the selected images from the CASIA dataset and compared them with [23] and [24] to analyze the performance of the proposed method.

The method of [23] is based on a normalized gray level co-occurrence matrix on 8x8 DCT coefficients and, using the Bayesian posterior probability map, localized the tampering objects. Whereas, the method [24] uses a deep learning method based on Fully Convolutional Networks (FCN) with Region Proposal Network (RPN) to localize the tampered region. To evaluate the superiority of the proposed method, we compared our results with conventional and deep learning methods.

Table I contains the Comparative results of the proposed method with [23] and [24] methods on both datasets based on average F-measure. FCN methods [24] prove to have superior performance than the conventional statistical-based methods [23]. From the results, it is evident that BAG noise on individual objects in the proposed method enables us to have much superior performance than [23].

The method is robust when it has a stable performance even after applying some post-processing operations on the spliced image. To evaluate the proposed method's robustness, we applied JPEG compression with different quality factors, Gaussian blur, and added Gaussian noise to all the spliced images and tested.

For JPEG compression, eight different quality factors ranging from 20 to 90 are considered. For Gaussian blur, Gaussian smoothing kernel with standard deviation $\sigma = 1.0$ is used, and for Gaussian noise, the variance of 0.03 and 0.05 are considered.

The evaluation results on IM Dataset has been shown in Table II. As the quality factor (QF) in JPEG compression decreases and additional post-processing operations included, the FCN and NGLCM methods decrease in their average F-measure values. In contrast, the proposed method has superior as well as stable performance even in such situations.

The IM dataset images are very high-resolution, and we try to downscale the quality factor to the lowest level 20. Fig. 5 is a graph showing the proposed method's performance with other existing methods. Both FCN+RPN and NGLCM methods decreased their average F-measure as the JPEG compression quality factory is reduced towards 20. The proposed method outperforms and gives stable performance even when the quality factor reduces because the BAGs are affected only in those objects than the rest of the image.

B. Computational Complexity

The effectiveness of any method depends on its average computation time spent is minimal to get the desired result. In the proposed method, after segmenting the individual objects, we obtain BAG features from each object instead of the whole image, thereby saving a lot of computation time. For localization, also we used a simple statistical method instead of unsupervised learning techniques. Table III gives the average running time spent by each method. Among the methods, the proposed method takes less time than other methods.

IV. Conclusion

This paper is proposed an efficient method for splicing localization based on block artificial grids in a double compressed JPEG image. When a JPEG image spliced with another image's object, the block artificial grids move from 8x8 gridlines to its centre. Taking this clue, we exposed splicing forgery through object segmentation. The method is straightforward, effective than other conventional methods that use JPEG fingerprints. The proposed method also robust even when the quality factor is low in high-resolution JPEG compression. The method fails on low-resolution images, and we considered it as our future work.

References

[1] A. M. Qureshi and M. Deriche, "A review on copy-move image forgery detection techniques," in *IEEE 11th International Multi-Conference on Systems*, 2014, pp. 1–5.

[2] J. A. Redi, W. Taktak, and J. Dugelay, "Digital image forensics: a booklet for beginners," *Multimed Tools and Applications*, vol. 51, pp. 133–162, 2011.

[3] P. More, T. N. Shankar, and P. Borse, "Storage covert channel concealment in tcp field," *International Journal of Control Theory and Applications*, vol. 10(1), pp. 1–7, 2017.

[4] H. Farid, "Image forgery detection a survey," *IEEE Signal Processing Magazine*, vol. 26(2), pp. 16–25, 2009.

[5] G. K. Birajdar and V. H. Mankar, "Digital image forgery detection using passive techniques," *Digital Investigation: The International Journal of Digital Forensics & Incident Response*, vol. 10(3), pp. 226–245, 2013.

[6] P. Kakar, N. Sudha, and S. W., "Exposing digital image forgeries by detecting discrepancies in motion blur," *IEEE Transactions on Multimedia*, vol. 13, pp. 443–452, 2011.

[7] T. N. Shankar and K. Spurthy, "Intrusion detection system using frequent item set in manet," *Journal of Adv Research in Dynamical and Control Systems*, vol. 10(1), pp. 356–362, 2018.

[8] T. N. Shankar, R. K. Senapati, P. M. K. Prasad, and G. Swain, "Volumetric medical image compression using 3d listless embedded block partitioning," 2016, vol. 1(2100), pp. 1–16.

[9] N. T. Babu, "Fpga implementation of hybrid system using timing attack resistant cryptographic technique," *Advances And Applications In Mathematical Sciences*, vol. 17(1), pp. 271–292, 2017.

TABLE I. COMPARATIVE RESULTS ON CASIA AND IMD DATASETS USING AVERAGE F-MEASURE

Method	CASIA 2.0	IMD
FCN+RPN	0.7388	0.6234
NGLCM	0.6524	0.5572
Proposed	0.7852	.0692

TABLE II. COMPARATIVE RESULTS FOR ROBUSTNESS ON IM DATASET USING AVERAGE F-MEASURE

Method	(JPEG Compression)		(Gaussian Blur)	(Gaussian Noise)	
	QF=50	QF=70	$\sigma = 1.0$	variance=0.03	variance=0.05
FCN+RPN	0.4365	0.6158	0.6187	0.6132	0.6092
NGLCM	0.3934	0.4323	0.5412	0.5389	0.5395
Proposed	0.6418	0.6596	0.7520	0.7514	0.7520

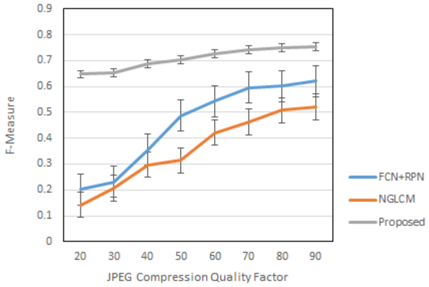

Fig. 5. Comparative Results of JPEG Quality Factory with F-Measure

TABLE III. AVERAGE RUNNING TIME

Method	FCN+RPN	NGLCM	Proposed
(Average Running Time in sec)	97.3	78.9	16.8

[10] P. N. R. L. C. Sekhar and T. N. Shankar, “Review on image splicing forgery detection,” *International Journal of Computer Science and Information Security*, vol. 14(11), pp. 471–475, 2016.

[11] K. Spurthy, T. N. Shankar, and R. K. Senapati, “Improving authentication of an iris recognition system by digital signature via elliptic curve cryptosystem,” 2016.

[12] K. Bahrami, A. C. Kot, and L. Li, “Blurred image splicing localization by exposing blur type inconsistency’,” *IEEE Trans. Inf. Forensics Security*, vol. 10(5), pp. 999–1009, 2015.

[13] Y. Zhang, C. Zhao, Y. Pi, and L. S., “Revealing image splicing forgery using local binary patterns of dct coefficients,” in *Liang Q. et al. (eds) Communications, Signal Processing, and Systems. Lecture Notes in Electrical Engineering*. Springer, 2012, vol. 202), pp. 181–189.

[14] K. Spurthy and T. N. Shankar, “An efficient cluster-based approach to thwart wormhole attack in adhoc networks,” *International Journal of Advanced Computer Science and Applications*, vol. 11(9), pp. 312–316, 2020.

[15] Z. He, W. Lu, W. Sun, and J. Huang, “Digital image splicing detection based on markov features in dct and dwt domain,” *IEEE Transactions on Pattern Recognition*, vol. 45(12), pp. 4292–4299, 2012.

[16] K. L. P. Rao, K. R. Rao, and K. R. R. M. Rao, “Adaptive energy efficient decentralized hierarchical dynamic cluster based routing protocol in wsn,” *Advances And Applications In Mathematical Sciences*, vol. 17(1), pp. 185–191, 2017.

[17] T. Bianchi and A. Piva, “Image forgery localization via block-grained analysis of jpeg artifacts,” *IEEE Transactions on Information Forensics and Security*, vol. 7(3), pp. 1003–1017, 2012.

[18] M. Zampoglou, S. Papadopoulos, and Y. Kompatsiaris, “Large-scale evaluation of splicing localization algorithms for web images,” *Multimedia Tools and Applications*, vol. 76(4), pp. 4801–4834, 2017.

[19] Z. Lin, J. He, X. Tang, and T. Ck, “Fast, automatic and fine-grained tampered jpeg image detection via dct coefficient analysis,” *Pattern Recognition*, vol. 42, pp. 2492–2501, 2009.

[20] I. Amerini, R. Becarelli, R. Caldelli, and A. Del Mastio, “Splicing forgeries localization through the use of first digit features,” *IEEE International Workshop on Information Forensics and Security (WIFS)*, pp. 143–148, 2014.

[21] H. Farid, “Exposing digital forgeries from jpeg ghosts,” *IEEE Trans. Inf. Forensics Security*, vol. 4(1), pp. 154–160, 2009.

[22] T. Bianchi, A. De Rosa, and A. Piva, “Improved dct coefficient analysis for forgery localization in jpeg images,” in *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, 2011, pp. 2444–2447.

[23] F. Xue, W. Lu, Z. Ye, and H. Liu, “Jpeg image tampering localization based on normalized gray level co-occurrence matrix,” *Multimedia Tools and Applications*, vol. 78, pp. 9895–9918, 2019.

[24] B. Chen, X. Qi, Y. Wang, Y. Zheng, H. J. Shim, and Y. Shi, “An improved splicing localization method by fully convolutional networks,” *IEEE Access*, vol. 6, pp. 69 472–69 480, 2018.

[25] Q. Wang and R. Zhang, “Double jpeg compression forensics based on a convolutional neural network,” *EURASIP Journal of Information Security*, vol. 2016(1), pp. 23–30, 2016.

[26] Y. Zhang, J. Goh, L. L. Win, and V. Thing, “Image region forgery detection: A deep learning approach,” *Proceedings of the Singapore Cyber-Security Conference (SG-CRC)*, vol. 14, pp. 1–11, 2016.

[27] B. Liu and C. M. Pun, “Locating splicing forgery by fully convolutional networks and conditional random field,” *Signal Process., Image Communication*, vol. 66, pp. 103–112, 2018.

[28] J. H. Bappy, C. Simons, L. Nataraj, B. S. Manjunath, and A. K. Roy Chowdhury, “Hybrid lstm and encoder-decoder architecture for detection of image forgeries,” *IEEE Transactions on Image Processing*, vol. 28(7), pp. 3286–3300, 2018.

[29] K. He, G. Gkioxari, P. Dollar, and R. Girshick, “Mask r-cnn,” in *IEEE International Conference on Computer Vision (ICCV)*, 2017, pp. 2980–2988.

[30] W. Abdulla, *Mask r-CNN for object detection and instance segmentation on Keras and TensorFlow*. RCNN, 2017. [Online]. Available: https://github.com/matterport/Mask

[31] W. Li, Y. Yuan, and N. Yu, “Passive detection of doctored jpeg image via block artifact grid extraction,” *Signal Processing*, vol. 89, pp. 1821–1829, 2009.

[32] J. Dong, W. Wang, and T. Tan, “Casia image tampering detection evaluation database,” in *IEEE China Summit and International Conference on Signal and Information Processing, Beijing*, 2013, pp. 422–426.

[33] V. Christlein, C. Riess, J. Jordan, C. Riess, and E. Angelopoulou, “An evaluation of popular copy-move forgery detection approaches,” *IEEE Transactions on Information Forensics and Security*, vol. 7(6), pp. 1841–1854, 2012.