DNN Based Ant Colony Optimization for Video Tampering Detection

Siva Prasad Patnayakuni

Senior Data Engineer, HEB, Texas, USA

Summary
Lately, videotape tampering process becomes easier due to the rapid-fire advancements in stoner-friendly editing software and multimedia technology(e.g., Mokey by Imagineer Systems, and Photoshop and Premiere by Adobe). This technologies may largely tamper the original images, so that the followership gets mislead. First, an image is divided into blocks of different sizes by a rate-distortion-based modified horizontal-vertical partition scheme. Statistical redundancy of quantized DCT portions of each image block is reduced by a bit-plane dynamical arithmetical coding with a sophisticated environment modeling. The compressed frames are also segmented by introducing Watershed segmentation fashion. Then, the region- grounded approach where the target structure is regarded as a homogeneous region which is determined by a hunt process guided by applicable criteria for unity. This unity features are also handed to DNN(deep neural network) for phony discovery. In this videotape forensic process, DNN classifier is included for phony discovery. The CNN classifier is included in colorful being phony discovery ways. But, in our work we include DNN because it contains number of retired layers which give accurate results for this phony discovery process. To ameliorate the DNN performance, Ant Colony Optimization (ACO) algorithm is introduced in this phony discovery fashion. Every niche and corner of this world we can suitable to find the surveillance cameras for security purpose. But, some fraudsters perform phonies in this recorded videos for their own benefits. To identify this, a lot of phony discovery ways are coming into actuality. So in this work, we introduce the DNN grounded ACO to perform the phony discovery in images. This perpetration is reused in python simulation platform. The parametric evaluations are taken in terms of F1- Score, accuracy, Precision, Recall and the Experimental results will give bettered performance in videotape phony discovery.

Keywords:
DCT Compression, Least Mean square, watershed segmentation, Deep neural network, optimization.

1. Introduction

The digital videotape technology’s rapid-fire growth makes tampering or editing a videotape sequence veritably simpler than ahead. An object can be removed in a videotape sequence using dominant software for videotape editing like Apple Final Cut Pro, Adobe Premiere, and Adobe after goods. lately, to a lesser extent the exploration of videotape tampering discovery was riveted by numerous experimenters. Discovery of object phony is a new theme in the field of unresistant digital videotape forensic exploration(1- 2). A videotape comprises of a series of images called frames.

In the temporal, spatio-temporal, and spatial fields the videotape phony attacks are performed. Only in the both spatiotemporal and spatial disciplines the dupe- paste and Region splicing tampering arises and in the temporal field the Frame insertion, junking, copying and shuffling will do. It's spontaneous that also on vids the discovery process of image tampering can also be used(3). still suitable results as anticipated won't be produced owing to complex phases in vids similar as frequent moving objects or noise sustained as a result of contraction. also, banning the information of the temporal sphere causes heavy calculation cost. Into two types the videotape tampering recognition styles can be codified i.e. active and unresistant(similarly nominated as eyeless).

In active styles digital hand and watermarking comes, to validate the videotape certain important information ‟ s will be deliberately bedded into it. Tampering takes place if any change is made to the bedded information(4- 7). Only a many bias has the installation to bed a digital hand or a watermark to the captured videotape. Once tampering is done, in similar situations these styles conceivably fails before fitting digital hand or watermark. substantially in natural conditions certain information of the vids similar as metadata won't be available. Because of the intelligent models development the timber of undoubted fake videotape content has enlarged. For some times the picky revision of the image contend was done, but the use of same approaches to videotape demands exorbitantly labour regarding its mass use. In a videotape if every frame is considered as a separate image, there could be conceivably a lot of images to reuse effectively. These issues could be handled with better computing power, also with the DNNs elaboration. lately in numerous operations the Deep literacy styles have seen a huge success. So in colorful new fields these approach has been used, similar as identification of camera model(,9), steganalys is(10), image regain forensics(11), discovery of image manipulation(12), dupe- move phony discovery(CMFD) of image(13), and so on. Particularly the Generative inimical Networks(GANs) are applied for altering the original videotape to recreate facial expressions of mortal(14), change the meteorological conditions(15) and for face- switching operations(16).

There-enactment of mortal face is clearly a new still is a common exploration area where simply, the speaking

Manuscript received September 5, 2022
Manuscript revised September 20, 2022
https://doi.org/10.22937/IJCSNS.2022.22.9.108

head is altered visually to imitate the a alternate actors facial expression(14, 17, 18) or to brace with another audio track(19, 20). These are having some of the simple operations for case moviere-dubbing in a several language or producing new movie scenes with an iconic actor ‟ s old videotape videotapes, although fake contents can also be produced. In certain situations, to reliably wisecrack mortal eyes the fake content is surely enough. From(18) the authors realised that piecemeal from arbitrary guessing the performance of mortal observers is better when aiming to determine whether the footages of facialre-enactment was genuine or intertwined. still, the DNN conceivably can liberate fluently the authentic and forged footage. The main donation of this work is to identify the videotape phony . Stronger forensic attestations are handed by videotape sequences than the motionless images. therefore, surveillance videotape, as important substantiation, is frequently used in the case disquisition. The development of phony software ‟ s increase the threat of phony in videos. This phony is performed in an accurate manner, so it takes time to descry this phony and also set up precious. To minimize this cost and time of phony discovery, colorful exploration process are continuing in this field. In this watershed segmentation process is included to member the frames into number of regions. This uprooted features are veritably important precious for DNN to perform the phony discovery process. The association for this entire paper is In Section 2. Some being styles that are enforced for videotape phony discovery is bandied. The complete frame of this proposed system is bandied in Section. 3, then some details regarding DCT contraction, regularized LMS Algorithm for noise cancellation, watershed segmentation for point birth, Deep NN, and ACT are handed. The experimental analysis and outgrowth of this proposed system is bandied in Section. 4. Incipiently, the conclusion for the proposed system is handed in Section. 5.

2. Related Work

Y Yao et al., 2017[21] to recognize the item based fabrication performed overvideo, for that a profound learning-based framework was proposed. This approach applied convolutional brain organization (CNN) for separating highlights naturally from the patches of the info picture. The conventional CNN was presented in this strategy for imitation recognition. Prior to giving the information picture to CNN, process the whole video outlines by going it through three pre-handling layers. To diminish the fleeting overt repetitiveness among video outlines, an edge outright distinction layer was incorporated. To deal with the computational issues of picture convolution, a maximum pooling layer was utilized, and the lingering signal was upgraded by a high-pass channel layer.

Raveendra etal.(2020)(22) attained OK - tuned AlexNet with DWT- DCT markov features model to identify inter frame tampering discovery. For the discovery of particular object, Markov grounded fashion was used. The ImageNet dataset was applied to retraining the data and the discovery of inter frame videotape tampering achieved using complication neural(CN) networkmodel. This model attained the delicacy rate about99.16. In this system, the other results attained the values of 99.25,99.07 and99.166 for perfection, recall and F1- score, independently.

Durgaetal.(2018)(23) proposed short term hierarchical fast milepost model- grounded on videotape segmentation. In this system, a short term hierarchical fast milepost(SHFW) algorithm was introduced to the frames. The segmentation process was done grounded on the operation of the K- means clustering system. Using a fast milepost algorithm, the way for the frames to final parts was attained. In order to estimate the intensity between the edges, binary tree complex sea(DTCW) transfigure was used. The testing was done using videotape segmentation standard dataset that correspond of 60 test vids and 40 training vids. The performance of volume perfection(VP) recall and boundary perfection(BP) recall was estimated. This model achieved the average perfection of0.51 and0.48 for both VP and BP recall, independently.

Natarajan etal.(2021)(24) introduced a rapid-fire bilateral sludge process grounded smoothen edgepreservation.The captured image contains noise that can affect the analysis of better results. This model involved enhancing fashion to ameliorate the discrepancy and brilliance of the digital images. For the operation of edge filtering scheme, the quality of both image and videotape was bettered because of birth and preservation for minimizing the noise. Themulti-resolution processing was reduced significantly. To filter the discrepancy image edges, rapid-fire bilateral filtering fashion was involved. Using this system, the edge preservation of the dynamic discrepancy images were filtered. The labors of this filtering model had minimized by 54 related to other styles similar as stationary sea(SW) transfigure, Huber- markov arbitrary(HMR) field, and exact histogram(EH) specification.

Zampoglou et al(2019)(25) developed a multimedia forensics with deep literacy for detecting tampered vids. To change the homemade verification process, this model developed an automatic videotape manipulation discovery system was introduced. For the operation of the homemade styles, videotape filtering process was attained using mortal experts. In this model, two types of forensics pollutants was designed. They were separate cosine(DC) transfigure and image bracket. The DC transfigure was applied for videotape processing. Image bracket was attained with the help of deep convolutional neural(DCN) networks in order to minimize the videotapere-

quantization crimes. The dataset was attained from the NIST 2018 media forensics and the InVID fake videotape(FV) corpus. The delicacy and mean average performance for FV corpus were attained about0.6177 and0.6558, independently.
Johnston et al(2020)(26) proposed for analysing an authentic content grounded on videotape tampering localisationmethod.This system composed of complication neural network in order to estimate the quantization parameter along with inter/ intra pixel patches and deblocking sludge setting. To corroborate crucial frames, the decoded sequences of frame deltas can be deduced by convolutional neural network. Grounded on videotape tampering localisation fashion, it can be classified into three way similar as calculating frame deltas, relating crucial frames and localize tampering. The results of quantization parameter attained the values of7.18 for delicacy rate of convolutional neural networks.

Yang et al(2020)(27) enforced a videotape tampering discovery grounded depth face phony using convolutional neural network. With the reduction of phony in the face, deep complication neural(DCN) network used to assay the challenges involved in the information security. To gain better evaluation, this paper involved four different styles for the evaluation of a videotape grounded face phony . They were face position, discovery, scaling andinterception.The performance of face phony was detected using two different situations similar as videotape position and frame position. The dataset used for the trial purpose using Celeb- DF. In the frame position training process, the results of0.8256 and0.9016 for the delicacy and area under wind, independently. Grounded on the videotape position tampering discovery, the affair achieved the value of0.9548 for the area under wind.

Problem Statement

Currently, colorful problems are being in terms of making fake attestations using tampering the vids(or) images. For the identification of the verity, an advanced algorithm is essential to identify the tampered region in forged images(or) vids. Indeed however, numerous ways has been applied, because of the low rate of point matching, high calculation time and poor delicacy rate, being ways aren't suitable to effectively employed in videotape tamper discovery. Several experimenters are concentrated to prize the optimal features from the image data. Hence, those styles aren't suitable for the comparison of image region to descry the forged position. thus, the new proposed mongrel deepnet fashion is suitable for the discovery of tamper region of a videotape data.

3. Proposed Methodology

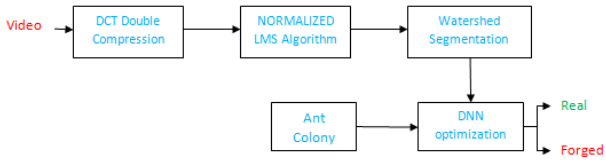


Fig.1 Proposed Method

3.1 DCT Compression

In Fig. 1, flowcharts of the proposed system are presented. As it can be seen from Fig. 2, the main difference between the proposed system and others consists in the following. 1) A modified horizontal-vertical (MHV) PS [28] is used. 2) A sophisticated bit-plane coding of the quantized DCT portions is exploited. 3) An effective system of de-blocking [29] is employed. The MHV scheme provides applicable adaption of the contraction together with simplicity of optimized partition. This, in turn, leads to the statistical unity of quantized DCT portions and simplicity of elimination of statistical redundancy of the quantized DCT portions by bit-plane coding. fresh gain on contraction performance is due to the junking of blocking vestiges. In this scheme, it's performed in the DCT sphere easing software and tackle consummations of the proposed system.

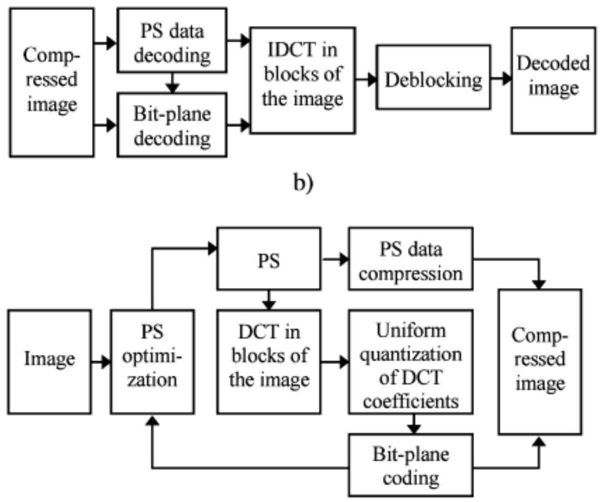


Fig. 2 Flowchart of a) Image compression. b) Image decompression.

3.2 Normalized LMS Algorithm

The traditional adaptive filtering trouble can be avowed in the subsequent way. Given an input signal u(n) and a asked signal d(n) determine the sludge, w, that minimizes the error, e(n), between the affair of the sludge, y(n), and the asked signal, d(n). For the case of Finite Impulse Responded(FIR) pollutants, an algorithm that solves this problem is the well known LMS. This is given by,

$$w(n\ 1) = w(n)\ \mu\ u(n) * e(n) \quad (1)$$

This equation updates the vector of the sludge portions w(n). The affair of the sludge is y(n) = $w^T(n)$ u(n) with u(n) = (u(n)... u(n − N1)) were N is the sludge length, and e(n) = d(n) − y(n). It's known that the LMS algorithm is only stable if the step size is limited, videlicet it should be equally commensurable to the power of the reference signal(1). This lead to the normalized LMS algorithm(NLMS), as represented in(2).

$$w(n\ 1) = w(n)\ \alpha\ u(n) * e(n)\ u(n)\ T\ u(n)^{*} \quad (2)$$

It's shown in(1) that this algorithm is stable as long as 0< α< 2 and of course, $u(n)^T u(n) * 6 = 0$. In order to help this last possibility, in practice, the algorithm is generally modified to,

$$w(n\ 1) = w(n)\ \alpha\ u(n) * e(n)\ u(n)\ T\ u(n) * q \quad (3)$$

were q is named to be small enough when compared with $u(n)^T u(n)^*$. This is generally chosen in an announcement croaker fashion. ways to select this value grounded on the proposed algorithm are presented in the paper.

3.3 Watershed Segmentation

Watershed segmentation is another region- grounded system that has its origins in fine morphology(Serra, 1982). The general conception was introduced by(Digabel and Lantuejoul, 1978). A break- through in connection was achieved by Vincent and Soille(1991) who presented an algorithm that's orders of bulks briskly and more accurate than former bones (see also(Hahn, 2005) for a discussion of the " watershed segmentation history "). Since also, it has been extensively applied to a variety of medical image segmentation tasks.

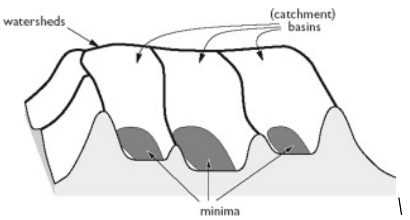


Fig 3. Principle of the watershed transform where the intensity values define hills and basins. (From: [Hahn, 2005]).

The controlled milepost system with morphological operation is robust and effective for early segmentation algorithms as reported by Refs.[30,31]. It's unnaturally different from conventional segmentation tools as it uses edge sensors to connect to pixel element extractors. The result to amend the under- and over-segmentation problems is by exercising milepost segmentation to divide images into unique regions grounded on their indigenous minima. For imbrication blood cell images, watershed segmentation is veritably effective with the use of a marker [32].

Watershed segmentation increases the architectural complication and computational rate of the segmentation algorithm. It also successfully overcomes the problems of high imbrication RBC.Fig.4 shows the pseudocode of the developed marker- controlled watershed system. The neighborhood of each single pixel is defined by the Euclidian distance dimension grounded on the periphery of the structuring rudiments so that Souter and wrongdoer return the sum of intensities of all pixels to the neighborhood of the external and inner indirect areas of the structuring rudiments, independently. Two counters are employed to count the pixels charted with corresponding structuring rudiments during each pass. The boxed area specifies the calculation of the intensity.

```
Algorithm 1 The pseudo-code of the algorithm
Input : f, Output : 1
v[p] ← 0, 1[p] ← 0, New_label ← 0, Scan_Step2 ← 1, Scan_Step3 ← 1 // Initialization
Scan from top left to bottom right : STEP1(p)
while Scan_Step2 = 1 do
    Scan image from top left to bottom right : STEP2(p)
    if v[p] is not changed then
        Scan_Step2 ← 0
    else
        Scan image from bottom right to top left : STEP2(P)
        if v[p] is not changed then
            Scan_Step2 ← 0
        end if
    end if
end while
while Scan_Step3 = 1 do
    Scan image from top left to bottom right : STEP3(p)
    if 1[p] is not changed then
        Scan_Step3 ← 0
    else
        Scan image from bottom right to top left : STEP3(p)
        if 1[p] is not changed then
            Scan_Step3 ← 0
        end if
    end if
end while
function STEP1(p)
    if v[p] ≠ 1 then
        for each n of p // n is neighbor pixel of p
        if f[n] < f(p) then v[p] ← 1
        end if
    end if
end function
```

Fig 4. The pseudocode of the algorithm.

The pseudocode of linked factors grounded on the marker-controlled watershed algorithm is shown in Fig.4, where p represents a pixel, I is the input preprocessed image(reused by sludge and morphological operation), and L is the segmented marker image.

I(p) defines the argentine position value of p, ne is the neighbor pixel of p, and I(ne) represents argentine position value of the coming separate pixel. The array d(p) is used to store the distance from the smallest pixels or table. still, the array l(p) is used to pasture the markers. DMAX and LMAX indicate the maximum distance and maximum value for the marker in the structure, independently. DMAX determines the space among the initial pixels of the initial row to the final pixel of the final row. Image scanning can be continued for step two if the v(p) array structure doesn't change. The image scanning can continue if l(p) array structure doesn't change.

3.4 DNN with Ant colony

The ACO grounded strategy works as follows. Given a potentially completely connected intermittent neural network – where each knot has a implicit connection to every knot in the posterior subcaste and to a separate knot in the intermittent subcaste – each connection between neurons can be seen as a implicit path for an ant(see Fig. 5). Every implicit connection is initialized with a base quantum of pheromone, and the master process stores the quantum of pheromone on each connection. Worker processes admit neural network designs generated by taking a named number of ants, and having them choose a path through the completely connected neural network poisoned by the quantum of pheromone on each connection.

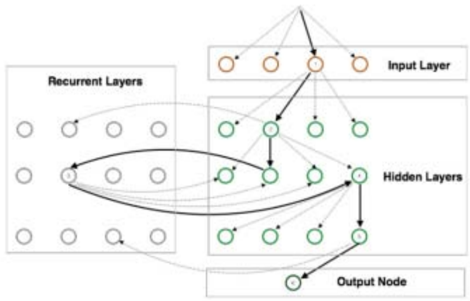


Fig. 5. Ants select a forward propagating path through neurons.

Evolving Deep intermittent Neural Networks Using Ant Colony Optimization Multiple ants can choose the connections between neurons. Those ant paths are be combined to construct a neural network design which is transferred to worker processes and trained on the input breakouts using backpropagation, evolutionary algorithms or any other neural network training algorithm. The master process maintains a population of the stylish neural network designs, and when a worker reports the delicacy of a recently trained neural network, if it improves the inhabitants, the master development will raise the pheromone on each relationship that was in to facilitate neural network. The master process periodically degrades pheromone situations, as is done in the standard ACO algorithm. This strategy allows the elaboration of intermittent neural networks with potentially numerous retired layers and hidden bumps, to determine what design can best prognosticate flight parameters(Fig. 6).

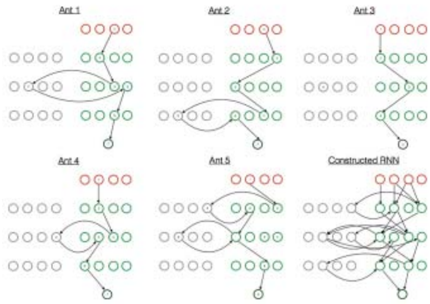


Fig. 6. The server creates neural networks for the workers.

4 Experimental Results

In this segment, different execution measurements are broke down to assess the successful exhibition of this proposed imitation discovery technique. The aftereffects of execution measurements like accuracy, review, precision, and F1-score are talked about in the accompanying subsections. The execution cycle for this interaction is performed on the Jupiter Notebook. The recordings for this phony recognition process is accumulated from YouTube. From Video altering dataset (VTD), we have taken the recordings for fabrication location, https://www.youtube.com/channel/UCZuuu-iyZvPptbIUHT9tMrA. rom these recordings, we have chosen 15 recordings for phony location. Here, 4370 edges are acquired from 15 distinct recordings. Among these 3059 are real and 1311 are fake. In our work, we utilized 70% of video outlines for training, 30% is for testing. Barely any pictures from VTD is displayed in Fig.7

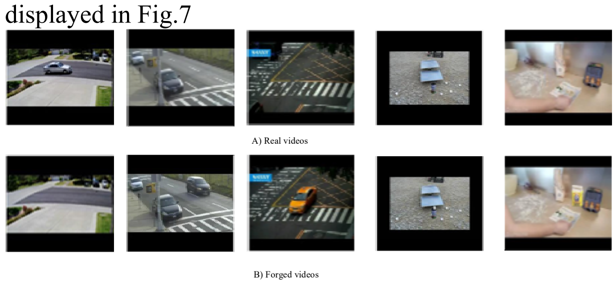


Fig. 7. Sample images from the VTD for A) Real and B) forged frames

The modernization boundaries utilized in the proposed system for falsification ID are: Maximum iteration is 100, number of neurons in the hidden layer (s2) is 600, learning rate (ε) is 1, penalty weight (β) is 0.4, probability (pr) is 0.2, and speed increase constants (C1,C2) is 2.05.

4.1. Evaluation metrics

Accuracy: The exactness aftereffects of this proposed approach is tracked down better compared to the current methods. The exactness aftereffect of this proposed technique is dissected for certain current methodologies. The recipe that is applied to distinguish this exactness esteem is displayed in the underneath condition (33),

$$A= (TP + TN)/(TP + TN + FP + FN) \times 100 \% \qquad (33)$$

Precision: The proportion got from the precisely recognized fashioned casing to the distinguished manufactured edge will decide the accuracy measurements. This accuracy is registered by:

$$P = TP/(TP + FP) \times 100 \% \qquad (34)$$

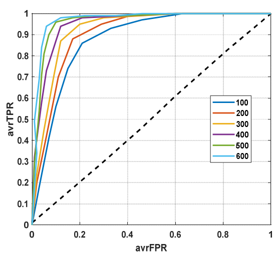
Recall: The fashioned parts not set in stone by review measurements. The proportion that exists among the precisely resolved fashioned edge to the all out outlines distinguished starting from the earliest stage of the manufactured picture decides the review. The recipe that is applied to ascertain the review esteem is planned in condition (35).

$$R = TP/(TP + FN) \times 100 \% \qquad (35)$$

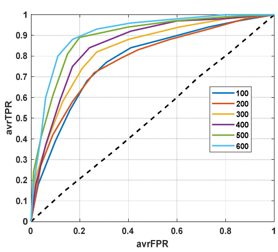
F1-score: The F1 score is likewise registered as an action that consolidates review and accuracy in a solitary worth. By applying accuracy and review, the F1 score esteem is distinguished. Condition (36), gives the specific F1 score esteem.

$$F1 = (2 \times P \times R)/(P + R) \times 100 \% \qquad (36)$$

The ROC for both compacted and uncompressed recordings with differing stowed away neurons is displayed in figure (8). The all out secret neurons utilized in the proposed DNN design is 600. For both packed and uncompressed recordings, the presentation altogether gets improved with expanding neuron numbers from 100 to 600. Subsequent to crossing as far as possible, the ROC bend portrays no improvement. The ROC of packed video is seen as obviously superior to the uncompressed video. This is on the grounds that the DNN engineering can learn more profoundly than the other ML calculations. Because of this explanation, the ROC of the uncompressed video likewise arrived at a high rate.



(a) ROC for compressed video



(b) ROC for uncompressed video

Fig. 8 ROC for a) Compressed video, b) Uncompressed video

In Table.1 shown, the assessment for this imitation recognition process is performed by utilizing four significant measurements like, TP (True positive), TN (True negative), FP (False positive), and FN (False negative). The quantity of produced outlines that are accurately perceived as fashioned is addressed as TP; comparably, the manufactured casings that are inaccurately perceived as genuine is addressed as FN. The genuine casings that are precisely identified as genuine is alluded as TN and the quantity of genuine edges that are mistakenly distinguished as fashioned is addressed by FP. The precision, accuracy, Fmeasure and recall measurements are assessed in this strategy to show its adequacy in phony recognition.

Table.1 Performance evaluation for proposed method

	Positive (%)	Negative (%)
True	95.04	93.75
False	5.54	4.97

In Table.2 shown, the proposed fabrication discovery strategy gives preferred acknowledgment results over the other four existing procedures they are Hong [33], Su[34], Sitara [35], and Abdalla[36]. The Accuracy, precision, recall and F-measure values accomplished by this proposed strategy is 94.72%, 94.49%, 95%, and 94.74%, individually.

Table. 2 Comparison of performance metrics of proposed method with two prevailing methods

Methods	A	P	R	F1-score
Proposed	**94.72**	**94.49**	**95**	**94.74**
Hong[33]	87.25	90.77	93.397	89.681
Su [34]	89.6	92.2	90.5	91.34
Sitara [35]	90.88	90.36	89.92	87.224
Abdalla[36]	91	69.63	80.42	88.35

5. Conclusion and future scope

The issue of Video Forgery is developing and should be battled to keep away from its enduring effect. The Forgery Detection procedures stay strange for recordings, and as a rule, the Image Forgery Detection methods have likewise been applied for something very similar. A considerable lot of the current methodologies incorporate just CNN for manufactured video recognition. Here, the DNN classifier is carried out alongside ACO, while this improvement calculation is remembered for this work to advance the weight boundary of DNN. This streamlining improves the presentation of DNN in imitation location. The accuracy, review, exactness, and F1-score of falsification identification are dissected to represent the adequacy of this proposed approach. The identification exactness accomplished by the proposed approach outflanks the other existing strategies. Lately, numerous falsifications are completed in video successions, yet at the same time, presently, a productive profound learning-based phony location for a wide range of imitations isn't created. A procedures might experience because of intricacy, less handling speed, besides some might neglect to recognize various sorts of frauds. So as future work, we intend to foster a calculation that actually recognizes a wide range of falsifications from recordings with less computational complexity, good Accuracy, and high speed.

References

[1] P. Johnstonand E. Elyan, "A review of digital video tampering: from simple editing to full synthesis." Digital Investigation (2019).

[2] Y. Yao, Y. Shi, S. Weng and B. Guan. "Deep learning for detection of object-based forgery in advanced video." Symmetry 10, no. 1, 2017, pp. 3.

[3] P. Johnston, E. Elyan and C. Jayne, "Video tampering localisation using features learned from authentic content." Neural Computing and Applications, 2019,pp.1-15.

[4] L. Bondi, S. Lameri, D. Güera, P. Bestagini, E. Delp and S. Tubaro, "Tampering detection and localization through clustering of camerabased CNN features." In 2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), 2017, pp. 1855-1864. IEEE. Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), 2017, pp. 1855-1864. IEEE.

[5] V. Joshi and S. Jain, "Tampering detection in digital video-a review of temporal fingerprints based techniques." In 2015 2nd International Conference on Computing for Sustainable Global Development (INDIACom), 2015, pp. 1121-1124. IEEE.

[6] T.M. Mohammed, J. Bunk, L. Nataraj, J.H. Bappy, A.Flenner, B. S. Manjunath, S.Chandrasekaran, A.K. Roy-Chowdhury and L.A. Peterson. "Boosting Image Forgery Detection using Resampling Features and Copy-move Analysis." Electronic Imaging 2018, no. 7 (2018),pp.1-7.

[7] L. Chen, X.Peng, J.Tian and J. Liu,"A learning-based approach for leaf detection in traffic surveillance video". Multidimensional Systems and Signal Processing, 2018, 29(4), pp.1895-1904.

[8] L.Bondi, L.Baroffio, D.Güera, P.Bestagini, E.J.Delp and S.Tubaro,"First steps toward camera model identification with convolutional neural networks". IEEE Signal Processing Letters, 2016, 24(3), pp.259-263.

[9] A.Tuama, F.Comby and M. Chaumont,"Camera model identification with the use of deep convolutional neural networks". In 2016 IEEE International workshop on information forensics and security (WIFS)2016, December, pp. 1-6. IEEE.

[10] G.Xu, H.Z. Wu and Y.Q. Shi,"Structural design of convolutional neural networks for steganalysis". IEEE Signal Processing Letters, 2016, 23(5), pp.708-712.

[11] P. Yang, R. Ni and Y.Zhao, "Recapture image forensics based on Laplacian convolutional neural networks". In International Workshop on Digital Watermarking,2016September, pp. 119-128. Springer, Cham.

[12] 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, 2016, June,pp. 5-10, ACM.

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

[14] S. Suwajanakorn, S. M. Seitz, I. Kemelmacher-Shlizerman, "Synthesizing obama: learning lip sync from audio", ACM Transactions on Graphics (TOG) 2017, 36 (4), 95.

[15] M.-Y. Liu, T. Breuel, J. Kautz, "Unsupervised image-to-image translation networks", in: Advances in Neural Information Processing Systems, 2017, pp. 700–708.

[16] H. Dong, P. Neekhara, C. Wu, Y. Guo, "Unsupervised image-toimage translation with generative adversarial networks", arXiv preprint arXiv:1701.02676.

[17] J. Thies, M. Zollh¨ofer, M. Stamminger, C. Theobalt, M. Nießner, "Face2face: Real-time face capture and reenactment of rgb videos", in: Computer Vision and Pattern Recognition (CVPR), 2016 IEEE Conference on, IEEE, 2016, pp. 2387–2395.

[18] A. R¨ossler, D. Cozzolino, L. Verdoliva, C. Riess, J. Thies, M. Nießner, "Faceforensics: A large-scale video dataset for forgery detection in human faces", arXiv preprint arXiv:1803.09179.

[19] T. Karras, T. Aila, S. Laine, A. Herva, J. Lehtinen, "Audio-driven facial animation by joint end-to-end learning of pose and emotion", ACM Trans actions on Graphics (TOG), 2017, 36 (4), 94.

[20] L. Chen, Z. Li, R. K Maddox, Z. Duan, C. Xu, "Lip movements generation at a glance", in: The European Conference on Computer Vision (ECCV), 2018.

[21] Y. Yao, Y. Shi, S. Weng and B. Guan, et al., Deep learning for detection of object-based forgery in advanced video. Symmetry, 10(1) (2017) 3.

[22] Raveendra, Malle, and K. Nagireddy. "Inter frame detection based on DWT-DCT Markov features and Fine-tuned AlexNet Model." *IJCSNS International journal of computer science and network security*, Vol.20 No.12, Dec 2020.

[23] Durga, R., G. Yamuna, and R. Barkavi. "Video Segmentation Using Short Term Hierarchical Fast Watershed Algorithm." In *2018 International Conference on Communication and Signal Processing (ICCSP)*, pp. 0281-0284. IEEE, 2018.

[24] Natarajan, Balakrishnan. "Smoothen Edge Preservation using Rapid Bilateral Filter Process." *Annals of the Romanian Society for Cell Biology* (2021): 11585-11590.

[25] Zampoglou, Markos, FoteiniMarkatopoulou, Gregoire Mercier, Despoina Touska, Evlampios Apostolidis, Symeon Papadopoulos, Roger Cozien, Ioannis Patras, Vasileios Mezaris, and Ioannis Kompatsiaris. "Detecting tampered videos with multimedia forensics and deep learning." In *International Conference on Multimedia Modeling*, pp. 374-386. Springer, Cham, 2019.

[26] Johnston, Pamela, EyadElyan, and Chrisina Jayne. "Video tampering localisation using features learned from authentic

content." *Neural computing and applications* 32, no. 16 (2020): 12243-12257.
[27] Yang, Tongfeng, Jian Wu, Lihua Liu, Xu Chang, and Guorui Feng. "VTD-Net: depth face forgery oriented video tampering detection based on convolutional neural network." In *2020 39th chinese control conference (CCC)*, pp. 7247-7251. IEEE, 2020.
[28] N. Ponomarenko, V. Lukin, K. Egiazarian, and J. Astola, “Modified horizontal vertical partition scheme for fractal image compression,” in Proc. 5th Nordic Signal Processing Symp., Hurtigruten, Norway, 2002.
[29] K. Egiazarian, J. Astola, M. Helsingius, and P. Kuosmanen, “Adaptive denoising and lossy compression of images in transform domain,” J. Electron. Imag., vol. 8, pp. 233–245, 1999.
[30] F. Tek, A. Dempster, I. Kale, Blood cell segmentation using minimum area watershed and circle radon transformations, Mathematical Morphology: 40 Years On, Proceedings of the 7th International Symposium on Mathematical Morphology, April 1820, 2005, pp. 441454.
[31] S. Kareem, R.C. Morling, I. Kale, A novel method to count the red blood cells in thin blood films, in: 2011 IEEE International Symposium on Circuits and Systems (ISCAS), IEEE, 2011.
[32] J.M. Sharif, et al., Red blood cell segmentation using masking and watershed algorithm: a preliminary study, in: 2012 International Conference on Biomedical Engineering (ICoBE), IEEE, 2012.
[33] J.H. Hong, Y. Yang and B.T. Oh, et al., Detection of frame deletion in HEVC-Coded video in the compressed domain. Digital Investigation, 30 (2019) 23-31.
[34] L. Su, T. Huang and J.Yang, “A video forgery detection algorithm based on compressive sensing”. Multimedia Tools and Applications, 2015, 74(17), pp.6641-6656.
[35] K. Sitara and B.M. Mehtre, et al., A comprehensive approach for exposing inter-frame video forgeries. In 2017 IEEE 13th International Colloquium on Signal Processing & its Applications (CSPA) IEEE (2017 March) 73-78.
[36] Abdalla, Y., Iqbal, M.T. and Shehata, M., 2019. Convolutional neural network for copy-move forgery detection. Symmetry, 11(10), p.1280.

Siva Prasad Patnayakuni is currently working as Senior Data Engineer in HEB, USA. His current research focus is on Distributed Data, Cloud Computing & Predictive Analytics and recommendations. He holds his Bachelors degree in Computer Sciences from Andhra University, India in 1996 and Masters in Computer Applications from Osmania University in 1999. He has been successfully designing and building large scale data warehousing, data pipelines, real-time analytics, reporting solutions for complex business problems. He created an intuitive architecture/frame works that helps organizations on effectively analyzing large scale of structured & unstructured data.