# Inter frame Tampering Detection based on DWT-DCT Markov Features and Fine tuned AlexNet Model

**Malle Raveendra[1]*, and K Nagireddy[2]**

[1]Ph.D Scholar, JNTUA, Anantapuramu, Andhra Pradesh,India
[2]Department of ECE, NBKRIST, Vidyanagar, Andhra Pradesh, India

**Summary**

Nowadays, videos are simply captured and viral in little time, and video editing has turn out to be more convenient with editing software. Therefore, the validity of the videos becomes more important. Inter frame video counterfeiting is the most common type of video spoofing method that is difficult to detect with the naked eye. So far, it has been suggested that some algorithms detect Inter frame counterfeits based on artisanal characteristics, but the accuracy and processing speed of these algorithms remain a challenge. This article proposes Markov based approach to detecting this particular object. First, the unique Markov characteristics in the DCT domain are extended to capture not only the inter-block correlation but also the intra-block association among the block DCT coefficients.after that, supplementary features are built in the DWT domain to distinguish three types of dependencies between the wavelet coefficients across positions, scales, and orientations. After that, we will introduce a video tampering detection method to detect Inter frame video tampering based on Convolutional Neural Network (CNN) models by retraining the accessible CNN model trained on the ImageNet dataset. The proposed method is based on state-of-the-art CNN models, which are retrained to exploit the spatio-temporal relations in a video to strongly detect Inter frame fakes and we have also proposed a confidence score instead of the score of raw output based on these networks, to increase the precision of the proposed method. Through the experiments, the detection Accuracy of the proposed method is 99.16%. This result has shown that the planned method has considerably accuracy and precision than other existing methods.

***Keywords:****Inter frame video fakes, Video manipulation, Artisan features, Markov-based approach, Convolutional Neural Network (CNN).*

## 1. INTRODUCTION

Today, smart phones, camcorders, and security cameras are extensively used in many areas of day to day. Especially in traffic lights, offices, houses, bedrooms and many other places that are monitored by cameras. In addition to that, video editing software like Video Editor, Adobe Photoshop, Window Movie Maker, and Adobe After Effect are readily accessible and easily used. These tools provide great support for editing video content without difficulty, even the edited content contrasts with the original content, leading to "seeing is not believing". Also, an authentic video gives stronger evidence than an authentic image in court. Hence, forensic video proves that video authenticity become an urgent prerequisite. So today forensic video has turn into a hot topic of awareness among investigators around the world.

Video criminology is the logical examination of a video for distinguishing and separating confirmations to check its genuineness, honesty or both. Video altering discovery is a subcategory of video legal sciences which looks at the video for content adjustments and may find the spatial or worldly areas of fabrication. These scientific methodologies can be either dynamic or detached dependent on the accessibility of earlier data about the video viable [1].

Dynamic strategies like computerized signature and watermarking [9-15] require pre-installed data in the record under scrutiny for its authenticity testing. The greater part of the video catching gadgets in market don't uphold this element. Additionally, it relies upon the sole carefulness of client if to insert this data. In this manner, these strategies are not solid. Latent or daze video altering identification strategies don't need earlier data for arranging a video as altered or not. These strategies are more dependable, all things considered, situations as they work by using the follows (or antiquities) of altering. Regardless of whether an endeavor is made to control the hints of altering, such endeavors will likewise bring about new alter follows [1]. This cycle of covering up or eliminating alter follows for misdirecting measurable investigation is called video hostile to criminology.

Video can be idea of as a grouping of pictures called outlines, shown throughout some stretch of time. Subsequently, the altering identification techniques created for picture crime scene investigation [16-20] can be applied at outline level. Time area, which is considered as

Manuscript received December 5, 2020.
Manuscript revised December 20, 2020.
https://doi.org/10.22937/IJCSNS.2020.20.12.1

the third element of video has critical function in video pressure. It might present movement relics and pressure commotion relying upon the video codecs utilized for pressure. Henceforth, the use of picture measurable procedures may neglect to catch video altering ancient rarities prompting bogus groupings.

Video frauds can be arranged into Inter edge and intra-outline imitations. In Inter edge imitations, outline in general goes through altering measure, while in intra-outline, outlines are controlled incompletely [21, 22]. In our work, we center around Inter edge video fabrications which can be ordered into outline erasure, outline addition, outline duplication and casing rearranging. Edge erasure manages occasions expulsion in video by eliminating the casings concerned. In edge inclusion, outlines duplicated from one video are stuck into another video. Edge duplication (or replication) includes duplicating of casings in a video and embeddings them at another fleeting areas of a similar video. Casing rearranging is another type of edge duplication where the replicated outlines are re-requested transiently before addition. Casing inclusion, outline duplication and casing rearranging can be utilized to fill the hole of erased outlines in a video.

In this study, we proposed a method that applies recent state-of-the-art CNN models, such as GoogleNet, ResNet, and VGG. These models were trained with more than one million images on ImageNet database [26], which were later fine-tuned and retrained on the target dataset for detecting some kinds of video Inter frame forgeries. We have also compared the efficiency of the models with each other to find out which architecture of the CNN model is suitable for detecting video Inter frame forgeries. In particular, the proposed models were not directly retrained from video frames, but they were retrained from the residual or optical flow between consecutive frames. We have performed many experiments to find out the best feature which was acquired for proposed methods. Besides, we have also conducted some tests to check the efficiency of transfer learning models trained on ImageNet database for this situation.

We have proposed a method for fine-tuning and retraining the state-of-the-art CNN models to detect video Inter frame forgeries. In addition, the confidence score is defined based on classification scores of the CNN model to enhance the effectiveness of the model and through many experiments, we have proven that the proposed method is efficient.

We have proposed four methods to build training datasets from original videos based on residual or optical flow features between adjacent/non-adjacent frames inside videos. Through experiments, we have suggested two methods that were most suitable to create datasets for training the state-of-the-art CNN models to detect video Inter frame forgeries.

The rest of the paper is planned as follows. Section 2 provides an overview on existing methods in the literature. Section 3 deals with background concepts required for this work. Section 4 discusses the proposed method for detecting various Inter frame forgeries. Experimental results and discussion are given in Section 5. Conclusions and future research directions are given.

## 2. RELATED WORK

Because of capacity and transmission imperatives, anxiety calculations are applied on recordings. For altering its substance, a video must be decompressed. Subsequent to performing imitation, it must be compacted once more. The works in [2-8] talked about techniques for grouping recordings as altered in the event that they have gone through twofold or numerous compressions. Wang and Farid used antiquities from the measurable circulation of P-outline expectation blunder grouping and Discrete Cosine Transform (DCT) coefficient conveyance of large-scale blocks (MB) of I outlines in [3] and [4] separately. Vazquez-Padin et al. [2] utilized the variety in expectation of MB sorts of P-outlines. Falsification follows in the forecast lingering succession got from movement vectors of closer views are utilized in [7].

In 2019, Raveendra et al. procedure is predicated on the consistency of rate field. With sequential system cancellation and structure duplication fraud tasks, some division pinnacles might be found in VFI grouping. what's more, thusly the summed up ESD investigate is applied to separate the pinnacles and decide the falsification sort in [54].

A technique for distinguishing moved I-outlines in twofold packed recordings utilizing convolutional neural organization is proposed in [29]. Twofold or numerous pressure discovery does not give an appropriate understanding into video alter location, as twofold compacted video may not be an altered one in all cases. A guide to help this contention, recordings showing up via web-based media stages go through additional pressure to conform to their guidelines. In such cases, as long as there is no adjustment in its substance, a video can't be named altered.

Melody et al. [30] proposed video record structure-based technique for video altering recognition. As it works with record signature, it can identify those altered recordings whose document marks are there in the put away information base. Thus, it bombs where a doctored video is made utilizing the culprit's own strategies than standard devices for altering.

The techniques in [42-49] used alter follows from the pressure space for Inter edge video fabrication location. Gironi et al. [42] expanded the work in [2] for distinguishing outline cancellation and edge addition. Edge

cancellation location utilizing movement repaid edge ancient rarity (MCEA) is examined in [39] and [44]. Su et al. [45] used intermittent ancient rarities in DCT coefficients of recompressed P and B outlines which emerges from outline move because of edge cancellation or edge addition. The techniques in [42-45, 48] are fruitful with fixed GOP structure recordings and bomb when a whole GOP or its products are competitors of falsification.

Casing duplication location techniques are proposed in [31-33]. Lin et al. [31] utilized connection of shading histogram contrasts between contiguous edges in competitor and question cuts for outline duplication recognition. Yang et al. [32] utilized relationship of Singular Value Decomposition (SVD) highlight of edge sub-groupings with that of the dubious casing sub-succession. Singh et al. [33] used connection and root mean squared mistake (RMSE) of mean and buildups in casing sub squares. These strategies bomb when copied outlines are rearranged prior to gluing.

Edge insertion recognition dependent on Multi-Level Subtraction (MLS) is proposed in [34]. It comprises of three degrees of deduction - pixel dim qualities from progressive edge sets are deducted in first level, the neighboring qualities got from first level deduction are deducted in second level and the qualities from second level are utilized for third level deduction. This strategy is touchy to pressure.

In Ref. [23], the model applied max-pooling method to the component maps. The model comprised of eight convolution layers, three pooling layers and one completely associated layer with a delicate max classifier. They applied the system on the public CASIA v1.0, CASIA v2.0 and DVMM datasets. The model utilized the SRM (spatial rich model) as a weight instatement rather than an irregular age. SRM assists with improving the speculation capacity and quicken the intermingling of the organization. Major SRM issues can be recorded as: it emerges overfitting sometimes, expanding the preparing time, and may different issues that lead the structure to undesirable outcomes. This system has another impediment is the corrected direct unit (ReLU) usage as an initiation work in the organization. ReLU units can be delicate during preparing and can "kick the bucket" which obviously gives disillusioning outcomes.

Jaiswal, A. et al. [24] proposed a system dependent on a blend of pre-prepared model resnet-50 and three discriminators (SVM, KNN, and Naïve Bayes). The model is applied and tried on CASIA V2.0 dataset [25]. The aftereffect of this calculation was not promising as the decision of resnet-50 was insufficient for the falsification issue. Resnet-50 development is mind boggling and it needs a huge handling time for playing out the cycle of both preparing and testing, and a major memory allotment which it is not acknowledged and legitimate in the genuine fabrication genuine critical thinking.

As this paper is enlivened by the AlexNet model engineering that was distributed and reported in 2012 [26], we looked and underscored the examination done on the recently distributed work that depends on the AlexNet model. It is valuable to specify that there are three exploration papers, a definitive found and known, which center their examination on AlexNet explicitly.

J. Ouyang et al. [27] proposed a system that can just distinguish duplicate move imitations utilizing AlexNet structure straightforwardly with no adjustments to the organization geography. They applied AlexNet on the ImageNet information base. They applied AlexNet model on UCID, OXFORD blossom, and CMFD datasets. The model got a decent exhibition to the falsification picture created consequently by PC with a basic picture duplicate move activity yet isn't powerful to the duplicate move phony picture of genuine situation. The outcome was not fulfilled enough and not hearty to duplicate move in a genuine situation.

They likewise demonstrated the idea that AlexNet can perform well in the phony discovery issue, and it was the principal execution of AlexNet in fraud location. This work was the motivation of different creators to begin taking a shot at AlexNet as pre-prepared organization design.

A. Doegar et al. [28] proposed AlexNet model-based profound with SVM classifier to be applied to the accessible benchmark dataset MICC-F220. The preparation was finished via preparing SVM utilizing AlexNet as profound highlights and for testing, the test pictures are applied to the prepared SVM to decide if the test picture is manufactured. This model structure yields incredible outcomes for the MICC-F220 dataset as it comprises of mathematical changes of a real image's. The exhibition of the profound highlights separated from the pre-prepared AlexNet based model is very agreeable, the best exactness of picture imitation discovery accomplished is 93.94%. This proposed strategy can just tackle the issue of duplicate move fabrications.

## 3. PROPOSED WORK:

The task of ordering video from a gathering of validated and controlled recordings is introduced as a two-class design acknowledgment issue. The particular highlights are caught by a Markov cycle in the DCT and DWT spaces. The proposed method utilizes a pre-named informational index to fabricate a computational model fit for identifying altering. It begins with highlight extraction to speak to every video in the dataset with a Frames include. At that point decrease the dimensionality of the component space and select the most applicable highlights to recognize pieces of information of changes because of control. Through regulated learning, an alexNet changed in

accordance with pronounce legitimate or altered is prepared and tried. The subtleties of these means are clarified in the accompanying subsections.

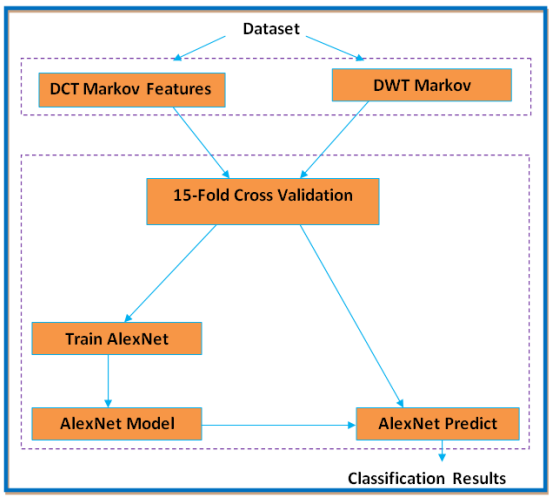



**FIGURE 1** Experimental procedure block diagram

### 3.1 Feature Extraction

A key issue in pattern recognition is feature extraction, which should provide a set of discriminatory features with low correlation with each other. For video tampering detection, the extracted features depend on the observation that tampering changes the correlation pattern between frames. In our case, we extract features from the space-time domain and merge them with features extracted from the DCT, DWT and DWT-DCT domain. In each domain, we model the statistical changes using a Markov process.

#### 3.1.1 Block DCT

The original Markov characteristics in the DCT domain proposed in [50]are very remarkable in capturing the differences between authentic and Tampering videos. They can be calculated by following the six steps below. First, apply the 8x8 Block Discrete Cosine Transform on the source image pixel matrix and the corresponding DCT coefficient matrix will be obtained. Second, round the DCT coefficients to a whole number and take an absolute value (denote the obtained matrix as M).

Thirdly, calculate the horizontal and vertical difference arrays using,

$$M_h(u,v) = M(u,v)-M(u+1,v) \quad (1)$$
$$M_v(u,v) = M(u,v)-M(u,v-1) \quad (2)$$

Fourth, enter a threshold T (T ∈ N +), if the value of an element in $\boldsymbol{M_h}$ (or $\boldsymbol{M_v}$) is greater than T or less than -T, replace it with T or -T, respectively. Here T is set to 4 (same thing below), to strike a balance between detection performance and computational complexity.

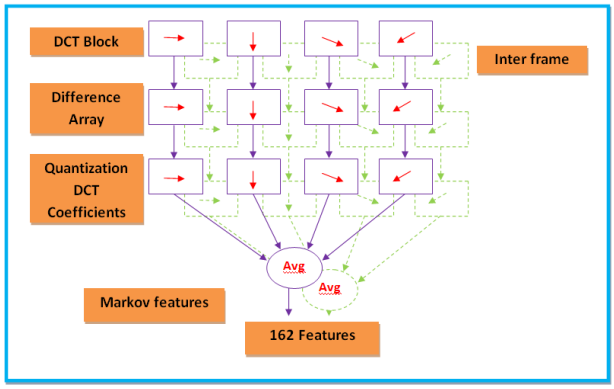



**FIGURE 2** Markov statistics based DCT Block

Fifth, compute the horizontal and vertical transition probability matrices of $\boldsymbol{M_h}$ and $\boldsymbol{M_v}$using,

$$P1h(i,j) = \frac{\sum_{u=1}^{su-2}\sum_{v=1}^{sv}\delta(Mh(u,v)=i,Mh(u+1,v)=j)}{\sum_{u=1}^{su-2}\sum_{v=1}^{sv}\delta(Mh(u,v)=i)} \quad (3)$$
$$P1v(i,j) = \frac{\sum_{u=1}^{su-1}\sum_{v=1}^{sv-1}\delta(Mh(u,v)=i,Mh(u,v+1)=j)}{\sum_{u=1}^{su-2}\sum_{v=1}^{sv}\delta(Mh(u,v)=i)} \quad (4)$$
$$P2h(i,j) = \frac{\sum_{u=1}^{su-1}\sum_{v=1}^{sv-1}\delta(Mv(u,v)=i,Mv(u+1,v)=j)}{\sum_{u=1}^{su-2}\sum_{v=1}^{sv}\delta(Mh(u,v)=i)} \quad (5)$$
$$P2v(i,j) = \frac{\sum_{u=1}^{su}\sum_{v=1}^{sv-2}\delta(Mh(u,v)=i,Mh(u,v+1)=j)}{\sum_{u=1}^{su-2}\sum_{v=1}^{sv}\delta(Mh(u,v)=i)} \quad (6)$$

where i, j ∈ {-T, -T + 1, ....., 0, ..... T-1, T}, Su and Sv denote the dimensions of the original source image. $\delta$ (.) = 1 if and only if its arguments are satisfied, otherwise $\delta$ (.) = 0.

Finally, all the elements of all the transition probability matrices are used as characteristics for the detection of image manipulation. the dimensionality of the final feature vector is (2T + 1) x (2T +1) x 4. As suggested above, through the original Markov features in the DCT domain mentioned above that are highlighted in capturing intra-block correlation, the correlation caused by the 8x8 blocking artifact is ignored. here we introduce the inter-block correlation between the corresponding coefficients. These additional Markov characteristics can be calculated similarly to the originals, such as

$$Nh(u,v) = M(u,v) - M(u+8,v) \quad (7)$$
$$Nv(u,v) = M(u,v) - M(u,v+8) \quad (8)$$

$$P3h(i,j) = \frac{\sum_{u=1}^{Su-16}\sum_{v=1}^{Sv}\delta(Nh(u,v)=i,Nh(u+8,v)=j)}{\sum_{u=1}^{Su-16}\sum_{v=1}^{Sv}\delta(Nh(u,v)=i)} \quad (9)$$

$$P3v(i,j) = \frac{\sum_{u=1}^{Su-8}\sum_{v=1}^{Sv-8}\delta(Nh(u,v)=i,Nh(u,v+8)=j)}{\sum_{u=1}^{Su-8}\sum_{v=1}^{Sv-8}\delta(Nh(u,v)=i)} \quad (10)$$

$$P4h(i,j) = \frac{\sum_{u=1}^{Su-8}\sum_{v=1}^{Sv-8}\delta(Nv(u,v)=i,Nv(u+8,v)=j)}{\sum_{u=1}^{Su-8}\sum_{v=1}^{Sv-8}\delta(Nv(u,v)=i)} \quad (11)$$

$$P4v(i,j) = \frac{\sum_{u=1}^{Su}\sum_{v=1}^{Sv-16}\delta(Nh(u,v)=i,Nh(u,v+8)=j)}{\sum_{u=1}^{Su}\sum_{v=1}^{Sv-16}\delta(Nh(u,v)=i)} \quad (12)$$

### 3.1.2 Block DWT

Wave analysis is useful for catching fleeting transient or confined change in signs. DWT is reasonable for video altering discovery. Numerous DWT-based strategies have been proposed before, for example, [35,36]. Nonetheless, most techniques manage all sub-groups freely after wave deterioration, paying little mind to the reliance between wave coefficients on positions, scales, and directions [37]. The methodology proposed in [38] portrays the three sorts of reliance between wavelet coefficients somewhat, yet the test results given by [39] show that it isn't appropriate for the recognition of misrepresentations of pictures. Here, they were grouped by the Markov Random Process (Transition Probability Matrix) to catch the previously mentioned three kinds of reliance between the wavelet coefficients and make the Markov attributes obtained in the DWT space a significant part in the whole picture graft location plot. Markov attributes in the DWT space can be determined as follows.

To start with, apply the 3-level Discrete Wave Transform on the pixel exhibit of the source picture, round all the friends of the got 12 sub-groups to the closest entire number, and take the outright worth. Mean the prepared estimation sub-groups, level detail sub-groups, vertical detail sub-groups, and corner to corner detail sub-groups as Ai, Hi, Vi, Di (I = 1,2,3), individually. Mean the pixel lattice of the source picture as A0 and view it as an estimate sub-band of level 0. Note that various wavelets work in an unexpected way, here we pick the discrete Meyer wavelet since it is symmetric and has the quality of minimal help in the recurrence area.

Second, think about one of the three sorts of reliance between wavelet coefficients: reliance through toxic substance. The utilization of change likelihood frameworks to speak to the reliance between the wavelet coefficients at the positions is very like portray the relationship between's neighboring DCT coefficients. Replacing F in Eqs. (1) and (2) with every one of the 13 wavelet sub-groups referenced over, 13 x 2 x 2 = 52 progress attributes can be gotten utilizing conditions (3) - (6). Third, think about another reliance between wavelet coefficients: the reliance between scales. Take the flat sub-groups Hi for instance, figure a framework like the distinction. Hence (2T + 1) x (2T +1) x 12 Markov Characteristics can be gotten. At long last, think about the last dependence between the wavelet coefficients: the reliance between directions. First ascertain the cross-contrast grids utilizing,

$$HV_i(x,y) = H_i(x,y) - V_i(x,y) \quad (13)$$
$$VD_i(x,y) = V_i(x,y) - D_i(x,y) \quad (14)$$
$$DH_i(x,y) = D_i(x,y) - H_i(x,y) \quad (15)$$

where $i \in \{1,2,3\}$. Then 3x3x2 = 18 more corresponding transition probability matrices, and therefore (2T + 1) x (2T + 1) x18 Markov characteristics can be obtained. In summary, we obtain (2T + 1) x (2T + 1) x 82 Markov characteristics in the DWT domain that characterize the three types of dependence between the wavelet coefficients. These Markov characteristics, together with those expanded in the DCT domain mentioned in Section 3.1.1, distinguish image manipulation from authentic ones.

### 3.1.3 Block DWT-DCT

In the combined approach, each unit of DCT and DWT area managed in the image in a sequence, one when it is the opposite, to compress the image much more and gain many higher compression ratios. Compression will increase with increasing window size for DCT and decrease with increasing window size for DWT. Then, in these 8x8 blocks, 2D-IDCT is performed followed by the first level 2D-IDWT in the 8x8 image block leading to 16 * 16 image blocks. Then the ordinal level 2D-IDWT is applied leading to 32x32 image blocks. Then merge is performed to retrieve the compressed image. The compressed image occupies less area compared to the initial image plus less than the area occupied by the image once compressed by DCT and DWT individually.

The hybrid DWT-DCT rework takes advantage of the properties of all DWT and DCT techniques and provides stronger compression. The input frame obtained from the video is first regenerated in 32x32 blocks. Each block is then reworked privately. The 32 × 32 block is converted to 16 × 16 once at a dwt level and discarding all coefficients except LL (that is, LH, HH, and HL). The second level of the two-dimensional dwt applies to the held LL coefficients. And this produces the Associate in Nursing 8 × 8 block once discarding all ICSH, HH, metric displacement unit coefficients, and LL-only protector. DCT is applied in this block. Lossy compression occurs once the transformation by DCT, the division is applied to the DCT coefficients that rounds the high frequency components to zero. The reverse, initial reverse division technique is completed and then the IDCT per 8 × 8 block is performed. Then the first level IDWT provides 16 × 16 blocks, and in addition, the second level of IDWT provides the 32 × 32 block. This technique is applied to the entire

image.

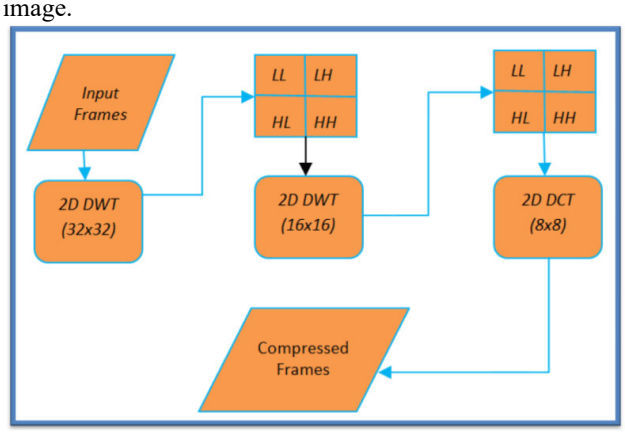



**FIGURE 3** Flow chart of DWT-DCT Block

### 3.1.4 Huffman Coding

The entropy encoding utilized here is Huffman encoding. This is frequently a lossless pressure strategy that appoints a prefix code called a Huffman code to the two info signals. The major arrangement of Huffman coding is to dole out each flexibly picture letters in order a specific assortment of pieces that doesn't surpass the memory limit. The length of the applied pieces relies upon the measure of data contained in the flexibly image. So the main arrangement of Huffman coding is to trade every one of the gracefully images with a simpler one and it is overseen bit by bit. This progression is kept to one side with just two images for the best code.

### 3.1.5 Contribution of different parts.

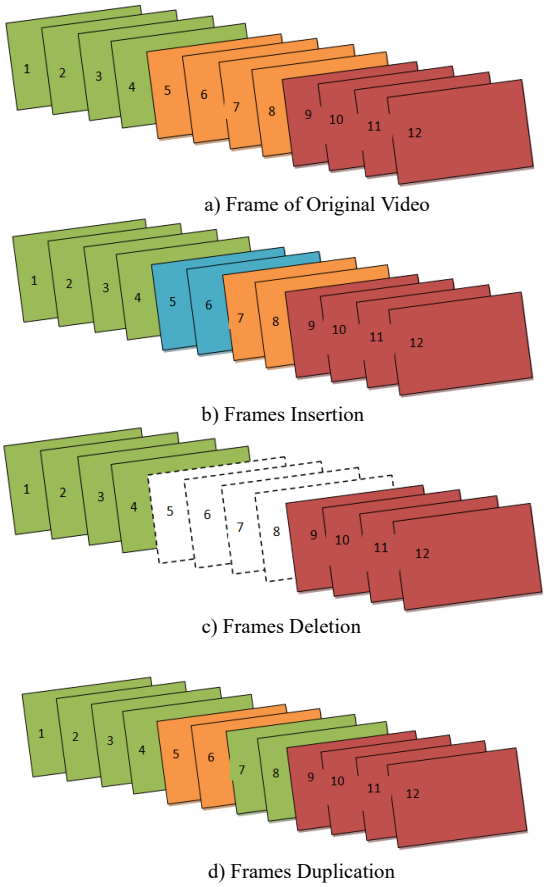
Since the proposed normal picture model comprises of two various types of Markov highlights, for example extended Markov includes in DCT space and Markov includes in DWT area, a few trials are additionally directed to analyze their separate commitments to the discovery execution. The outcomes are given in Table 1.

**TABLE 1 Summary of Markov features with Different Thresholds**

| Markov Features | | DCT | DWT | DWT-DCT |
|---|---|---|---|---|
| T=3 | TPR | 0.9778 | 0.9643 | 0.9710 |
| | TNR | 0.9720 | 0.9705 | 0.9712 |
| | DA | 0.9901 | 0.9898 | 0.9899 |
| T=4 | TPR | 0.9796 | 0.9701 | 0.9748 |
| | TNR | 0.9787 | 0.9730 | 0.9758 |
| | DA | 0.9934 | 0.9899 | **0.9916** |
| T=8 | TPR | 0.9715 | 0.9635 | 0.9668 |
| | TNR | 0.9754 | 0.9749 | 0.9739 |
| | DA | 0.9902 | 0.9822 | 0.9831 |

As appeared in Table 1, the extended Markov includes in DCT area play out a litter in a way that is better than the Markov includes in DWT space, this is likely because of DCT's boss capacity in energy compaction, which brings about more little coefficients in the comparing contrast clusters and makes them all the more effectively described by the limited progress likelihood networks. Additionally, it can likewise be seen that, by joining these two sorts of Markov highlights, better location execution can be accomplished.

### 3.2 Tampering Detection with CNN




a) Frame of Original Video

b) Frames Insertion

c) Frames Deletion

d) Frames Duplication

**FIGURE 4** Video inter frame forgeries

Fakes between video frames can include three types of fakes shown in Fig. 4 as follows:

a) Frames from the original video shown in fig. 4a, 1st to 12th frames with solid edge.

b) Inserting a sequence of frames shown in Fig. 4b; In Fig. 4b, a 5th to 6th frame sequence copied from a different video and then pasted after the 5th frame with dropout. This fake is often used to add events from a different video

to the video.
c) Removing a sequence of frames shown in Fig. 4c, the 5th to 8th frames with the dotted edge were removed to hide events within a video.
d) Duplication of a sequence of frames shown in Fig. 4d; In Fig. 3d, frames 3 and 4 were copied and then pasted at frames 7 and 8 in the same video without erasing the frame. This forgery is generally used to duplicate events in a video.

All the above video fakes can easily manipulate videos with one of the video content editing software like Adobe Photoshop, Adobe After Effect, Video Editor, and Window Movie Maker, etc. And those counterfeit videos would feature fingerprints, which are inconsistencies in the dimension of time. Pixel values between two consecutive frames in the manipulated position shown in Fig. 4. Those inconsistent pixel values are difficult to detect because they can generally be very small inconsistencies when manipulating videos in a sophisticated way. To detect those fingerprints, we have proposed a method by applying the powerful next-generation CNN models that train with AlexNet and adjust and then retrain them on the target dataset to detect those fingerprints. The detail of the proposed method is presented in the next section.

### 3.3 Methods for creating training datasets

For training a model to detect video fakes who need a large number of videos, including original videos and fake videos. To overcome the dearth of large video data sets for training models and to take advantage of previously trained models in the ImageNet database, we have proposed four methods to construct four different training data sets based on the residual or optical flow of adjacent or non-adjacent. stills in original videos. The residual or optical flow of adjacent frames in the original video has consistency, which is used to create negative samples. Otherwise, the residual or optical flow of non-adjacent frames in the original video has an inconsistency, which is used to create positive samples. In particular, these four methods create four data sets to retrain the model as follows: a) residual from two adjacent or non-adjacent frames, b) three gray value residuals on four adjacent or non-adjacent frames, c) optical flow of two adjacent or non-adjacent frames, and d) three optical flow magnitudes in four adjacent or non-adjacent frames. The details of each method are as follows:

Let $X = \{x^t\}$ is an original video.

Where, $t \in [1, T]$, T is the number of frames of the video. xt is the tth frame of the video.

**A. To Create a training dataset from the residuals of two adjacent or non-adjacent frames: Dataset1**

From the videos in the original video dataset, negative samples were created by subtracting two adjacent frames, and positive samples were created by subtracting two non-adjacent frames, particularly in the following steps:

***To create negative samples:***

Create $R = \{r^t\}$,
negative samples as follows:

```
 for t = 1: T-1 do
          r^t = x (t + 1) - x^t;
finish;
```

// for Where, $x^t$ is the tth frame of the video. $r^t$ is a remainder of two adjacent frames as the difference between two adjacent frames, considered as a negative sample.

***To create positive samples:***

Create $R' = \{r''^t\}$,
positive samples as follows:

```
for i = 1: T do
k = random (1: T);
/ * The distance between two non-adjacent frames is at least 15 frames. So, randomly generate k up to an absolute value of k greater than i-15 * /
while abs (k - i) <= 15 do
k = random (1: T);
finish;
  //While if k% 2 <> 0
r''^t = x^i - x^k more
r''^t= x^k - x^i
finish;
```

// Where, $r''^t$is a residual of two non-adjacent squares as the difference between two non-adjacent squares, considered a positive sample. In particular, in all experiments, we chose the distance between two non-adjacent frames within the video to be at least 15 frames because in reality Inter frame spoofing manipulations usually alter the length of the frame sequence by at least 15 frames. second.

**B. Create a training data set from three gray residuals - Dataset2**

***To creating negative samples:***

Create $R=\{r^t\}$, negative samples as follows:

```
for t = 1 : T-3 do
          r^t(:,:,1) = greyimage(x^(t+1)) - greyimage(x^t);
          r^t(:,:,2) = greyimage(x^(t+2)) - greyimage(x^(t+1));
          r^t(:,:,3) = greyimage(x^(t+3)) - greyimage( x^(t+2));
end; //for
```

Where, $x^t$ is the $t^{th}$ frame in the video.
$r^t$ is a sample including three residuals of grey values from four adjacent frames, considered a negative sample.

***To creating positive samples:***

Create R'={ $r'^{t}$ }, positive samples as follows: (5)

```
for i = 2 : T do
        k = random(1:T);
/* The distance between two non-adjacent frames is at least 15 frames. So, randomly generated k until absolute of k greater than i-15*/
while abs(k-i) <= 15 do
        k = random (1:T);
end; //while
if k%2 <> 0
        r''(:,:,1) = greyimage(x^i) - greyimage( x^(i-1));
        r''(:,:,2) = greyimage(x^k) - greyimage(x^i);
        r''(:,:,3) = greyimage(x^(k+1)) - greyimage(x^k);
else
        r''(:,:,1) = greyimage(x^k) - greyimage( x^(k-1));
        r''(:,:,2) = greyimage(x^i) - greyimage( x^k);
        r''(:,:,3) = greyimage(x^(i+1)) - greyimage( x^i)
end; //if
end; // for
```

Where, $r'^{t}$ is a sample that includes three residual values of gray in four non-adjacent frames considered as a positive sample. Similarly, the distance between two non-adjacent frames within the video is at least 15 frames.

C. Creation of a training data set from the optical flow of two adjacent or non-adjacent frames - Dataset3 Creating Dataset3 from the optical flow of two adjacent or non-adjacent frames is similar to creating Dataset1 by changing the residuals to the optical flow of two adjacent or non-adjacent frames.

D. Creating a training data set from three optical magnitudes - Dataset4 Creating Dataset4 from three optical flow magnitudes in four adjacent or non-adjacent frames is similar to creating Dataset2 by changing the gray values to the optical flow magnitude of four adjacent or non-adjacent frames.

# 4. EVALUATION

In this section, we introduce how to tune the next-generation CNN models, setting up to retrain the models in the target dataset, the results of testing the models, and preparing the data that we have collected and built for use in our experiments. . In addition to that, we have also compared the results with some more recent research that was done on the same data set.

## 4.1 Dataset description

Due to the scarcity of a large Inter frame forgery dataset to train the proposed CNN models, we have compiled a dataset with 300 original videos from five surveillance cameras from the VFDD dataset [41] that was taken from surveillance cameras in real life by our laboratory. This dataset was captured with various environments, such as inside and outside of schools, offices, dormitories, streets and buildings with different lighting conditions, during the day and at night with light and without light. The average length of the videos is 10 seconds.

To create the training dataset, we randomly selected 270 videos from this dataset and followed the dataset creation steps in section 3.3 to build four training datasets. Finally, we got four data sets, each of which has around 120,000 negative and positive samples.

To create a dataset for the test, we have used 25 original videos remaining from the dataset of 300 original videos. We manipulate these videos manually in different ways. In that way, we have 100 videos including 25 original videos and 75 counterfeit videos, including the three types of counterfeits between previous video frames. An important note here is that all the faked videos we have tampered with are not easily detectable with the naked eye. This entire dataset is published online at [40].

## 4.2 Fine-tuning and retraining models

To retrain the models in the target data sets, we have fine-tuned the leading edge models by removing the last three layers of those networks. Because the last three layers contain information on how to combine the characteristics that the network extracts into original class probabilities and labels. Then add three new layers to the layer chart, including a fully connected layer, a soft-Max layer, and a sort output layer. We have also configured the fully connected end layer to be the same size as the number of classes in the target dataset (this case is 2). To learn faster in the new layers than in the transferred layers, we have set the learning rate of the fully connected layer equal to 5. In addition, the rest of the training options were set as follows: 85% randomly selected from training data set for retraining, 15% for validation. We use the SGD optimization method which has a momentum contribution from the previous step of 0.85. The initial learning rate is 0.001, and the learning rate would drop 0.1 after 10 epochs; mini_batch_size is 10, max_epochs is 20 and is shuffled in each epoch, the L2 regularization is 0.0001.

## 4.3 Performance metrics

For testing, each video in the test dataset portion will be followed by the negative sample creation steps in section 3.3 From that, we would have a set of samples from each video. This set of samples is classified using the previous trained models. Finally, it is concluded that each video is a fake or an original

**TABLE 2 The Results ofproposed method that applies recent state-of-the-art CNN models**

| Fine tuned and retrained Model | TP(%) | TN(%) | FP(%) | FN(%) | P(%) | R(%) | DA(%) | F1(%) |
|---|---|---|---|---|---|---|---|---|
| ResNet18 | 98.12 | 98.73 | 3.55 | 2.93 | 96.51 | 97.10 | 96.81 | 96.80 |
| ResNet50 | 99.15 | 98.92 | 1.74 | 2.65 | 98.27 | 97.39 | 97.83 | 97.83 |
| VGG16 | 97.64 | 96.23 | 4.26 | 5.93 | 95.82 | 94.27 | 95.01 | 95.04 |
| GoogleNet | 96.61 | 95.76 | 5.31 | 8.32 | 94.79 | 92.07 | 93.38 | 93.41 |

video, which depends on the maximum of fcon values of that video. The video is original if max ($fcon(i)$) <Threshold, where $i\in$ 1: $T$ - 1, otherwise faked. And in all experiments, we set the threshold equal to 0.5. For the performance measure, we rely on the following criteria: The measures are used in this document as follows: True Positive (TP): falsified video declared falsified; False positive (FP): original video declared false; True Negative (TN): original video declared genuine; Sensitivity or True Positive Rate (TPR); False Positive Rate (FPR) and Detection Accuracy (DA) as follow:

**TABLE 3 Performance evaluation for proposed method**

| | Positive (%) | Negative (%) |
|---|---|---|
| TRUE | 99.89 | 99.73 |
| FALSE | 0.75 | 0.93 |

### 4.3 Performance metrics

**Detection Accuracy:**

$$DA = \frac{TP + TN}{TP + FP + TN + FN} \times 100\% \quad (16)$$

**Precision:**

$$P = \frac{TP}{TP + FP} \quad (17)$$

**Recall:**

$$R = \frac{TP}{TP + FN} \quad (18)$$

**F1 score:**

$$F1 = \frac{2 \times P \times R}{P + R} \quad (19)$$

### 4.4 Experiments & Results

The results in Table 2 show that typically the more model parameters, the greater the accuracy in detecting Inter frame video forgeries. But other than that, there are exceptions that the models based on Resnet50 and Resnet18 gave quite good results, amounting to 97.83% and 96.81% respectively, while the number of parameters is not very large. Their accuracies are on same level with some other models with more parameters, such as VGG16 and GoogleNet. Thus, the proposed method based on ResNet50 or ResNet18 may be suitable for use in the detection of Inter frame video in situations that require high processing speed or low hardware.

**TABLE 5 The results of a model trained from four datasets built from four different characteristics**

| Datasets | DA (%) | FPR (%) | TPR(%) |
|---|---|---|---|
| Dataset1 | 97.20 | 2.67 | 99.28 |
| Dataset2 | 85.62 | 14.25 | 93.32 |
| Dataset3 | 83.97 | 18.35 | 92.52 |
| Dataset4 | 95.11 | 4.35 | 97.36 |
| Dataset1 and Dataset4 | 99.165 | 0.75 | 99.89 |

The results of a model trained from four datasets built from four different characteristics are shown in Table 5 From these results, we found that features such as the residuals of two adjacent or non-adjacent frames and three amounts of optical flow on four adjacent or non-adjacent frames are usable. for training and classification in the proposed model. In particular, accuracy would be significantly increased to 99.165% by combining the two.

**TABLE 4 Comparison of proposed method with other prevailing Methods**

| Method | DA(%) | P(%) | R(%) | F1(%) |
|---|---|---|---|---|
| Proposed | 99.165 | 99.25 | 99.07 | 99.166 |
| Li[51] | 95.1 | 94.47 | 95.89 | 95.2 |
| Su[52] | 92.6 | 91.2 | 90.49 | 92.78 |
| Yu wang[53] | 80 | 88 | 81 | 84 |

Due to the lack of large databases, we conducted several experiments to compare model performance between models trained on the basis of learning transfer and foundation. It is worth noting that the accuracy of ResNet50 increased from 85.17% (when training from uncompressed video) to 97.83%, and ResNet18 increased from 83.23% (when training from uncompressed video) to 96.81% when training with Compressed video.

The results in Table 4 show that the proposed method of detecting Inter frame video forgeries has an accuracy of 99.165%, which is much better than the latest methods. It has been proven that the proposed method is significantly effective in detecting Inter frame video

forgeries.

## 5. CONCLUSION

Nowadays, with the dynamically developing hardware industry, and especially with the development of cameras that were used to measure everywhere such as traffic, home, school, office etc. Moreover, most people use smart phones that are also equipped with cameras. This allows videos to be recorded anywhere, manipulated at any time, and spread quickly across the Internet. Authentic video has great value as evidence. But so far, while there are some methods to authenticate videos, they are either inefficient or very slow. In this study, we proposed a method based on the most modern CNN models to detect Inter frame video forgeries which showed good and probable results, the accuracy is 96.81%, 97.83% and 99.165%. It was proved through experiments that the proposed method achieved much higher efficiency than the newest methods on the same data set.

In the future, we will conduct in-depth research to propose a suitable CNN architecture with fewer parameters and complexity to detect and classify different types of video fraud.

## 6. REFERENCES

[1] K. Sitara, B. M. Mehtre, Digital video tampering detection: An overview of passive techniques, Digital Investigation 18 (2016) 8–22.

[2] D. Vazquez-Padin, M. Fontani, T. Bianchi, P. Comesan˜a, A. Piva, M. Barni, Detection of video double encoding with GOP size estimation, in: 2012 IEEE International Workshop on Information Forensics and Security (WIFS), IEEE, 2012, pp. 151–156.

[3] W. Wang, H. Farid, Exposing digital forgeries in video by detecting double MPEG compression, in: Proceedings of the 8th workshop on Multimedia and security, ACM, 2006, pp. 37–47.

[4] W. Wang, H. Farid, Exposing digital forgeries in video by detecting double quantization, in: Proceedings of the 11th ACM workshop on Multimedia and security, ACM, 2009, pp. 39–48.

[5] J. Xu, Y. Su, Q. Liu, Detection of double MPEG-2 compression based on distributions of DCT coefficients, International Journal of Pattern Recognition and Artificial Intelligence 27 (01) (2013) 1354001.

[6] X. Jiang, W. Wang, T. Sun, Y. Q. Shi, S. Wang, Detection of double compression in MPEG-4 videos based on markov statistics, IEEE Signal processing letters 20 (5) (2013) 447–450.

[7] P. He, X. Jiang, T. Sun, S. Wang, Double compression detection based on local motion vector field analysis in static-background videos, Journal of Visual Communication and Image Representation 35 (2016) 55 – 66.doi:http://dx.doi.org/10.1016/j.jvcir.2015.11.014.

[8] A. Subramanyam, S. Emmanuel, Pixel estimation based video forgery detection, Signal Processing, IEEE, 2013, pp. 3038–3042.

[9] P. Rasti, S. Samiei, M. Agoyi, S. Escalera, G. Anbarjafari, Robust nonblind color video watermarking using QR decomposition and entropy analysis, Journal of Visual Communication and Image Representation 38 (2016) 838 – 847. doi: http://dx.doi.org/10.1016/j.jvcir.2016.05.001.

[10] N. Sahu, A. Sur, SIFT based video watermarking resistant to temporal scaling, Journal of Visual Communication and Image Representation 45 (2017) 77 – 86. doi: https://doi.org/10.1016/ j.jvcir.2017.02.013.

[11] Y. Tew, K. Wong, R. C.-W. Phan, K. N. Ngan, Multi-layer authentication scheme for HEVC video based on embedded statistics, Journal of Visual Communication and Image Representation 40 (2016) 502 – 515. doi: http://dx.doi.org/10.1016/j.jvcir.2016.07.017.

[12] M. Fallahpour, S. Shirmohammadi, M. Semsarzadeh, J. Zhao, Tampering detection in compressed digital video using watermarking, IEEE Transactions on Instrumentation and Measurement 63 (5) (2014) 1057–1072. doi:10.1109/TIM.2014.2299371.

[13] P. K. Mishra, I. Hooda, Robust adaptive watermarking in video for protecting intellectual properties, in: 2016 3rd International Conference on Computing for Sustainable Global Development (INDIACom), 2016, pp. 3128–3131.

[14] A. M. Buhari, H. C. Ling, V. M. Baskaran, K. Wong, Low complexity watermarking scheme for scalable video coding, in: 2016 IEEE International Conference on Consumer Electronics-Taiwan (ICCE-TW), 2016, pp. 1–2. doi:10.1109/ICCE-TW.2016.7520710.

[15] O. S. Faragallah, Efficient video watermarking based on singular value decomposition in the discrete wavelet transform domain, AEU - International Journal of Electronics and Communications 67 (3) (2013) 189 – 196. doi:http://doi.org/10.1016/j.aeue.2012.07.010.

[16] L. Verdoliva, D. Cozzolino, G. Poggi, A feature-based approach for image tampering detection and localization, in: 2014 IEEE International Workshop on Information

Forensics and Security (WIFS), 2014, pp. 149–154. doi:10.1109/WIFS.2014.7084319.

[17] C. Li, Q. Ma, L. Xiao, M. Li, A. Zhang, Image splicing detection based on markov features in QDCT domain, Neurocomputing 228 (2017) 29 – 36, advanced Intelligent Computing: Theory and Applications. doi: http://doi.org/10.1016/j.neucom.2016.04.068.

[18] X. Zhao, S. Wang, S. Li, J. Li, Passive image-splicing detection by a 2- D noncausal markov model, IEEE Transactions on Circuits and Systems for Video Technology 25 (2) (2015) 185–199. doi: 10.1109/ TCSVT.2014. 2347513.

[19] V. T. Manu, B. M. Mehtre, Detection of Copy-Move Forgery in Images Using Segmentation and SURF, Springer International Publishing, Cham, 2016, pp. 645–654. doi:10.1007/978-3-319-28658- 7_55.

[20] C.-M. Pun, B. Liu, X.-C. Yuan, Multi-scale noise estimation for image splicing forgery detection, Journal of Visual Communication and Image Representation 38 (2016) 195 – 206. doi: http://dx.doi.org/10.1016/j.jvcir.2016.03.005.

[21] R. D. Singh, N. Aggarwal, Detection and localization of copy-paste forgeries in digital videos, Forensic Science International 281 (Supplement C) (2017) 75 – 91. doi: https://doi.org/10.1016/ j.forsciint.2017.10.028.

[22] S. Chen, S. Tan, B. Li, J. Huang, Automatic detection of object-based forgery in advanced video, IEEE Transactions on Circuits and Systems for Video Technology 26 (11) (2016) 2138–2151. doi:10.1109/TCSVT.2015. 2473436.

[23] Rao, Y.; Ni, J. A deep learning approach to detection of splicing and copy-move forgeries in images.In Proceedings of the 2016 IEEE International Workshop on Information Forensics and Security (WIFS), AbuDhabi, UAE, 4–7 December 2016; pp. 1–6.

[24] Jaiswal, A.K.; Srivastava, R. Image Splicing Detection using Deep Residual Network. SSRN Electron. J. 2019,8, 102.

[25] Dong, J.; Wang, W. CASIA v1.0 and CASIA v2.0 Image Splicing Dataset. Available online: https://www.kaggle.com/sophatvathana/casia-dataset(accessed on 28 September 2019).

[26] Krizhevsky, A.; Sutskever, I.; Hinton, G.E. Pdf ImageNet classification with deep convolutional neural networks. Commun. ACM 2017, 60, 84–90.

[27] Ouyang, J.; Liu, Y.; Liao, M. Copy-move forgery detection based on deep learning. In Proceedings of the 2017 10th International Congress on Image and Signal Processing, BioMedicalEngineering and Informatics (CISP-BMEI), Shanghai, China, 14–16 October 2017; pp. 1–5.

[28] Doegara, A.; Duttaa, M.; Kumar, G. CNN based Image Forgery Detection using pre-trained AlexNetModel.Proc. Int. Conf. Comput. Intell. IoT (ICCIIoT) 2019, 2.

[29] P. He, X. Jiang, T. Sun, S. Wang, B. Li, Y. Dong, Frame-wise detection of relocated i-frames in double compressed H.264 videos based on convolutional neural network, Journal of Visual Communication and Image Representation 48 (2017) 149 – 158. doi:https://doi.org/10.1016/j. jvcir.2017.06.010.

[30] J. Song, K. Lee, W. Y. Lee, H. Lee, Integrity verification of the ordered data structures in manipulated video content, Digital Investigation 18 (2016) 1– 7. doi: http://doi.org/10.1016/j.diin.2016.06.001.

[31] G.-S. Lin, J.-F. Chang, C.-H. Chuang, Detecting frame duplication based on spatial and temporal analyses, in: Computer Science & Education (ICCSE), 2011 6th International Conference on, IEEE, 2011, pp. 1396–1399.

[32] J. Yang, T. Huang, L. Su, Using similarity analysis to detect frame duplication forgery in videos, Multimedia Tools and Applications 75 (4) (2016) 1793–1811.

[33] V. K. Singh, P. Pant, R. C. Tripathi, Detection of frame duplication type of forgery in digital video using sub-block based features, in: International Conference on Digital Forensics and Cyber Crime, Springer, 2015, pp. 29– 38.

[34] C. C. Huang, Y. Zhang, V. L. L. Thing, Inter frame video forgery detection based on multi-level subtraction approach for realistic video forensic applications, in: 2017 IEEE 2nd International Conference on Signal and Image Processing (ICSIP), 2017, pp. 20–24. doi:10.1109/SIPROCESS.2017. 8124498.

[35] W.Chen,Y.Q.Shi,W.Su,Imagesplicingdetectionusing 2-Dphase congruency and statistical moments of characteristic function, in:Imaging:Security, Steganography,and Watermarking of Multimedia Contents,2007, p. 65050R.

[36] W. Lu, W. Sun, F.-L. Chung, H. Lu, Revealing digital fakery using multi- resolution decomposition and higher order statistics, Engineering Applications of Artificial Intelligence 24 (4) (2011) 666–672.

[37] A.Srivastava,A.Lee,E.Simoncelli,S.-C.Zhu, On advances in statistical modeling of natural images, Journal of Mathematical Imaging and Vision18(1) (2003)17–33.

[38] H.Farid,S.Lyu,Higher-order wavelet statistics and their application to digital forensics, in:International Conference on Computer Vision andPattern Recognition Workshop, Madison, WI,USA, 2003, pp.1–8.

[39] T.-T. Ng, S.-F. Chang, Blind Detection of Digital Photomontage using Higher Order Statistics, Technical Report 201-2004-1, Columbia University, 2004.

[40] Xuan Hau, N. and H. Jongjian, VIFFD - The data set for detecting video Inter frame forgeries. Mendeley Data; http://dx.doi.org/10.17632/r3ss3v53sj.4, 2019. V4.

[41] Al Hamidi, S., VFDD (Video Forgery Detection Database) Version 1.0. http://sites.scut.edu.cn/misip/main.psp, 2017.

[42] A. Gironi, M. Fontani, T. Bianchi, A. Piva, M. Barni, A video forensic technique for detecting frame deletion and insertion, in: 2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), IEEE, 2014, pp. 6226–6230.

[43] Y. Su, J. Zhang, J. Liu, Exposing digital video forgery by detecting motion- compensated edge artifact, in: Computational Intelligence and Software Engineering, 2009. CiSE 2009. International Conference on, IEEE, 2009, pp. 1–4.

[44] Q. Dong, G. Yang, N. Zhu, A MCEA based passive forensics scheme for detecting frame-based video tampering, Digital Investigation 9 (2) (2012) 151–159.

[45] Y. Su, W. Nie, C. Zhang, A frame tampering detection algorithm for MPEG videos, in: Information Technology and Artificial Intelligence Conference (ITAIC), 2011 6th IEEE Joint International, Vol. 2, IEEE, 2011, pp. 461– 464.

[46] M. C. Stamm, W. S. Lin, K. J. R. Liu, Temporal forensics and antiforensics for motion compensated video, IEEE Transactions on Information Forensics and Security 7 (4) (2012) 1315–1329. doi:10.1109/TIFS.2012.2205568.

[47] T. Shanableh, Detection of frame deletion for digital video forensics, Digital Investigation 10 (4) (2013) 350–360.

[48] X. Kang, J. Liu, H. Liu, Z. J. Wang, Forensics and counter anti-forensics of video Inter frame forgery, Multimedia Tools and Applications 75 (21) (2016) 13833–13853. doi:10.1007/s11042-015-2762-7.

[49] L. Yu, H. Wang, Q. Han, X. Niu, S. Yiu, J. Fang, Z. Wang, Exposing frame deletion by detecting abrupt changes in video streams, Neurocomputing 205 (Supplement C) (2016) 84 – 91.doi: https://doi.org/10.1016/j.neucom.2016.03.051.

[50] Raveendra M, Nagireddy K. “DNN based moth search optimization for video forgery detection”. International Journal of Engineering and Advanced Technology, 9(1), 1190-1199, 2019.

[51] Li, Q., Wang, R. and Xu, D., 2018. An Inter frame Forgery Detection Algorithm for Surveillance Video. Information, 9(12), p.301.

[52] Su, L. and Li, C., 2018. A novel passive forgery detection algorithm for video region duplication. Multidimensional Systems and Signal Processing, 29(3), pp.1173-1190.

[53] Liyang Yu, Huanran Wang, Qi Han, XiamuNiu, S.M. Yiu,Junbin Fang and Zhifang Wang, Exposing Frame Deletion by Detecting Abrupt Changes in Video Streams, Neurocomputing,http://dx.doi.org/10.1016/j.neucom.2016.03.051.

[54] Raveendra M, Nagireddy K. "An Uncovering Video Between Edge Phony Based On Velocity Field Reliability", International Journal of Science & Engineering Development Research (www.ijrti.org), ISSN:2455-2631, Vol.4, Issue 8, page no.61 - 66, August-2019.

## AUTHORS PROFILE

**MalleRaveendra** born in 1983 in a remote village in Andhra Pradesh, INDIA and completed B.Tech from S.V University in the year 2006 and obtained M.Tech from Bharath University in the year 2009. Presently he is Pursuing Ph.D. in Jawaharlal Nehru Technological University, Anantapuramu, Andhra Pradesh, INDIA. His areas of interest include Image and video processing.

**Dr.K.Nagi Reddy** born in 1974 in a remote village in Andhra Pradesh, INDIA and completed AMIETE in the year 1996, obtained M.Tech from JNT University in the year 2001 and obtained a Ph.D. from S.VUniversity. Presently he is working as Professor in NBKR. Institute of Science & Technology, Vidyanagar, Nellore (dt), Andhra Pradesh, INDIA. He is a life member of ISTE, IETE. His areas of interest include Image and Video Processing.