International Journal of Current Engineering and Technology
E-ISSN 2277 – 4106, P-ISSN 2347 – 5161

Available at http://inpressco.com/category/ijcet

Research Article

Forensic Video/Image Analytics using Deep Learning and Enhancement Techniques

Yuvraj B. Deshmukh and Dr. S. K. Korde

Department of Computer Engineering, Pravara Rural Engineering College, Loni, India.

Received 10 Nov 2020, Accepted 10 Dec 2020, Available online 01 Feb 2021, **Special Issue-8 (Feb 2021)**

Abstract

Forensic video/Image analysis is used in a many visual digital evidence analysis scenarios. Video/Image forensics is important & necessary to show that images and videos which are used as potential evidence in court of law are verifiably true. The popularity of smart digital/mobile devices and due to increasing number the low cost of surveillance systems, several forms of visual data are widely being used in digital forensic investigation. Digital videos are many times used as key evidence sources in evidence identification, analysis, presentation, and report. The motto here is to develop advanced forensic video analysis methods & techniques in order to perform effective digital forensic investigation. Further the aim is to develop a forensic video/image analysis framework that employs an efficient video/image processing techniques using deep learning, such as object detection framework YOLO V3, Head pose estimation technique in wild and face detection, color, face and skin detection methodology, also enhancing algorithm for the low quality of video footage analysis which consists of adaptive video enhancement algorithm based on Contrast Adaptive Histogram Equalization (CLAHE) technique, Contrast Exposure Fusion Algorithm, Dynamic Histogram Equalization (DHE) which are introduced to improve the quality of visual data for the use of digital forensic identification & investigation. The technique of Video/Image tampering detection using state-of-art techniques and textual enhancement are also incorporated in order to assist in truthful analysis of visual data evidence. The framework would deploy recent techniques and algorithms which will assist examiner to perform visual analysis of evidence data.

Keywords: *Video/Image Analysis, deep learning, object detection, Enhancement Algorithm, head pose estimation, tampering detection.*

Introduction

The field of computer vision is emerging rapidly and has many more computational processing power, memory & digital resources today involved. Video today has become the most popular media for communication and entertainment. Whereas automatic analysis and understanding the content of a video through technology is one of the important goals of computer vision. One of the very basic problems is to model & detect the appearance and behavior of several objects in a video. Such kind models mainly & mostly depends on the problem definition. Today more than 1 Billion new users visit YouTube each month, while watching 6 billion hours of videos, and are uploading 100 hours of video for each minute. Cameras today are hence termed to be ubiquitous and sifting through this ocean of data, which pose a major global challenge. This widespread visual digital media data, in recent years has made the automated analysis of its content necessary in several various fields of application. Videos may be analyzed for the tracking, detection and recognition of human activities in a variety of applications, ranging from indoors environments to outdoors locations, such as videos in the wild. As the subject being widespread and visual data is being used in several aspects of life, content analysis of such visual data is termed to be important.

The proposed framework is divided into three parts. The first part consists, of deep learning based video analysis methodologies such as YOLO object detection technique and Head Pose Estimation in-the-wild using Perspective-n-Point and dlib face detector. The second part consists of video tampering detection using inter-intra frame analysis techniques, Video/Image Summary and Hashing. Whereas the third part consists of image enhancement module including Contrast Limited Adaptive Histogram Technique (CLAHE), Dynamic Histogram Equalization, Contrast Exposure Fusion technique and wiener filter textual enhancement, color pattern detection modules.

Literature Survey

The paper "Video-Based Evidence Analysis and Extraction in Digital Forensic Investigation"[1] explain

about many digital forensic image/ video analysis techniques including deep learning object detection framework using YOLO technique, CLAHE image enhancement technique, color and shape based object detection methodologies. The paper also addresses several forensic visual data analysis problems and solutions in order to perform digital video based analysis in forensic environment. The several latest techniques of visual data analysis have been elaborated, also several experimental results of object detection and image enhancement techniques are displayed which conclude that YOLO as object detection can be used to detect the several crime objects and suspects and establish link between several crime scene objects and suspects. The CLAHE enhancement technique is termed to be useful technique in enhancement of natural and artificial images in many low light and poor visual data scenarios. [1]

J. Redmon, " YOLOv3: An Incremental Improvement". [29] Authers present some updates to YOLO! They made a bunch of little design changes to make it better. The article compares the results of several multistage and single stage detectors of object detection. The results show that YOLO V3 of object detection outperforms other object detectors in terms of speed and accuracy.

Convolutional neural networks (CNNs) are termed to be excellent performance on the head pose estimation problem under controllable conditions, whereas their generalization ability in the wild needs to be improved.[4] To address this issue, authors propose an approach consisting of facial landmark information into the task simplifier and landmark heatmap generator which is constructed before the feed-forward neural network.[4] The technique uses information to normalize the face shape into a canonical shape in order to generate a landmark heatmap, which is based on the transformed facial landmarks to assist in feature extraction. This method can be used for enhancing generalization ability in the wild.[4] The method elaborated here was trained on 300W-LP and tested on AFLW2000-3D dataset.[4] The result shows increase in accuracy which improves from 88.5% to 99.0% and mean average error decreases from 5.94 to 1.46 on AFLW2000-3D.[4] The technique of facial landmark which is introduced so that the CNN could extract features that reflect head pose more efficiently, which significantly improves the accuracy of head pose estimation in the wild.[4]

Tampering detection is termed to involve set of several techniques which perform the task of identification of manipulated videos. In a tempered video frames of video can be inserted, deleted, duplicated or shuffled. To perform this task techniques such as optical flow, correlation between suspicious frames and consistency of velocity field intensity can be used. [17] There exists several techniques for detection of inter frame forgery. [18], [19], [20] Which uses optical flows to detect inter frame forgery, when there will be a discontinuity in the optical flow variance sequence in cases where the frames have been replaced, deleted or inserted. [21] The detection of inter frame forgery is done by analyzing velocity filed consistencies. [22] Even the technique of double MPEG compression can be used to detect frame manipulations. The Copy move and copy paste forgery has been detected by several techniques. These include block matching techniques and key point matching techniques [23] [24]. The technique of SIFT (Scale Invariant Feature Transform) features to match copied regions. Splicing is also addressed via different approaches. [25] All these techniques are mostly depend on the assumption that the spliced area will be different in terms of a several fundamental aspects. [26] These techniques uses DCT (Discrete Cosine Transform) coefficients to detect double quantization effect of JPEG compression. [27] Skin detection is the process of finding skin-colored pixels and regions in an image or a video. Applications of skin detection are personality recognition, human tracking, gesture analysis & adult content filtering and etc. [10] The techniques like HSV range color detector and backprojection can be used for this purpose.

The new technique of Image Contrast Enhancement using Exposure Fusion Framework is termed to be most important in enhancement of color properties of a dark image. [30]

The key challenge with Forensic Video/Image processing system is that, the analysis techniques involved in visual data analysis should not change/alter the original objects & subjects of visual data. In forensic video analysis, the quality of video is always an important challenge. The enhancement of the low quality of video is very important for assisting & conducting efficient forensic investigation.

Proposed Methodology

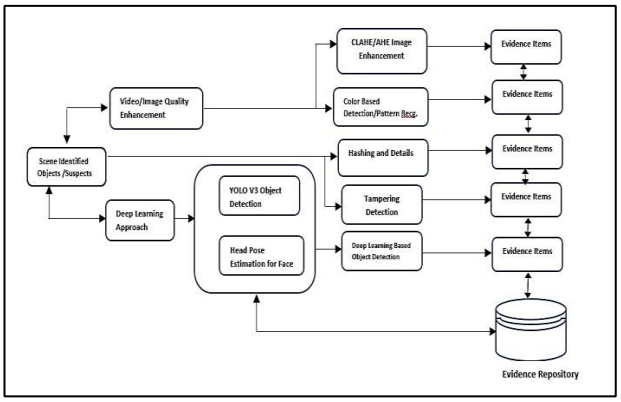

Fig.1. Framework for Forensic Video/Image Analysis System

In proposed framework, to achieve the goal of truthful visual data analysis and in order to adapt the new techniques of visual data analysis new modules are implemented which address the several problems of visual data analysis from forensic perspective. Forensic

visual data analysis today lacks adaption of new visual data analysis techniques such as deep learning and pattern recognition. The proposed framework can be used to solve the several problems of diversified nature of visual data analysis and could assist an examiner to perform truthful analysis of visual data.

Proposed System

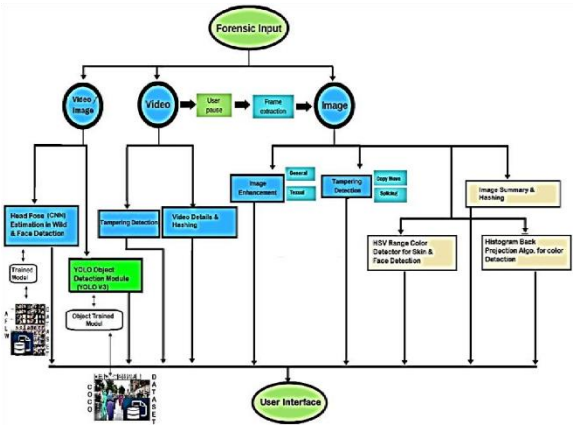

Fig. 2 Proposed System for Video/Image Analysis

We propose to develop truthful forensic Video/Image analysis system with proposed system and designing following modules for video/Image Analysis. The system is divided into several subparts involving deep learning based video/image analysis, tempering detection of visual data, color based analysis and image enhancement.

1) YOLO V3 Object Detection Module: It is an effective object detection framework using deep features. In our project the module performs the role of identification of different objects using state-of-art technique of object detection. The results are generated with greater accuracy and also the model is trained with large iterations of COCO dataset.

2) Head Pose Estimation using CNN Module: The module performs head pose estimation determining the pith, yaw and angle of the face and it also performs face detection in wild using CNN. The technique is based on perspective-n-point and dlib face detector algorithm. The technique is useful in many different real life scenarios of man machine interface. Here we use this technique to identify and detect suspects. The model is trained on Annotated facial landmarks in wild (AFLW) dataset.

3) Image Enhancement Module: This module consists of Contrast and Histogram Equalization Algorithms for Image Enhancement. These algorithms are used for enhancement of several color properties of image, this module is further divided into three different algorithmic image enhancement techniques such as i) Contrast Exposure Fusion Framework Algo, ii) Dynamic Histogram Equalization Algo.

iii) Contrast Limited Adaptive Histogram Equalization Algo. & Textual image enhancement based on Weiner filter is also included in order to eliminate noise and blur of a given image. In many situations of motion blur involved in visual data Weiner filter is termed to important technique in order to visualize blurred motion text.

4) Tampering Detection Module: The techniques of inter-intra frame tempering detection are used for video and image forgery detection. The technique such as inter frame forgery detection based on optimal flow technique is used for video tempering detection. Whereas techniques of intra frame detection are based on clustering and threshold, which are used towards detection of copy move and splicing type of image forgery. These techniques determine the truthfulness of visual data. Although these techniques are termed to be passive techniques of forgery detection towards visual data, these are termed to be authentic techniques of tampering detection of visual data. 5) ***Hashing and details Module:*** This module calculates hash value of image and video file, it also determines crucial metadata of a given file. Hash value function determine the authenticity and integrity of digital data and are termed to be digital fingerprint of digital data, also hash values are considered as crucial part of digital forensic analysis.

6) ***Color based Detection Module:*** This module uses techniques such as histogram backprojection, HSV range color detector in order to detect a certain color and process the image using color as fingerprint. These techniques are termed to be pattern recognition techniques for color based detection. These techniques are often used in tracking real time objects and suspects in many servellience applications.

B. Algorithms

1) Algorithm for object detection and tracking :[1]

Input: Video or camera input V
Output: Labeled V
1: v video_capture(V)
2: vg video_group(v)
3: for each frame fi 2 vg do
4: fhsv covColor(fi; COLOR_RGB2HSV)
5: fshsv split(fhsv)
6: for each channel c 2 fhue; saturatin; valueg do
7: threshld(fshsv[c]; dth[c]; low[c]; upper[c])
8: end for
9: fresult bitwise_and(dth[0]; dth[1]; dth[2])
10: end for
11: V update(v; vg)

2) Algorithm for video image enhancement using AHE (Adaptive Histogram Equalization)[1]

Input: Video or camera input V
Output: Labeled V
1: v video_capture(V)
2: vg video_group(v)
3: for each frame f 2 vg do
4: Initialize array Hist to zero;

5: for every contextual pixel j do
6: Hist[g(j)] D Histl [g[j] C 1
7: end for
8: CHistl D Pl kD0 Hist(k) 9: l0 D CHistl _ L=W2
10: vg update(f ; l0)
11: end for
12: V update(v; vg)

C. Flowcharts

1) Flow charts for Image Enhancement Techniques: i) Contrast Exposure Fusion Algorithm:

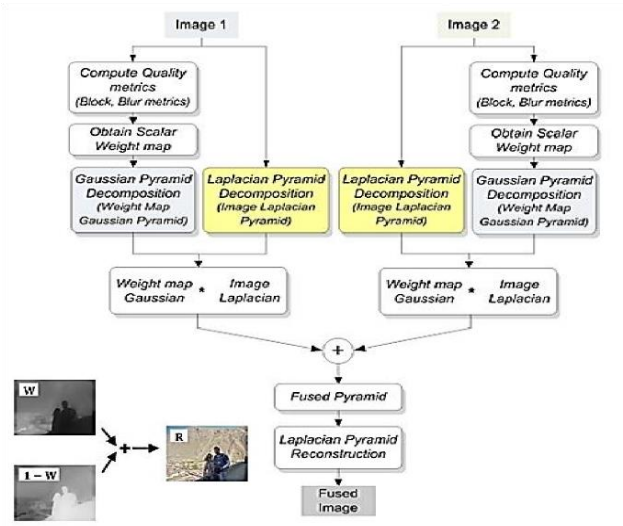

Fig.3. Flowchart for Contrast Exposure fusion Algo.[30]

ii) Contrast Limited Adaptive Histogram Equalization and Dynamic Histogram Equalization (DHE):

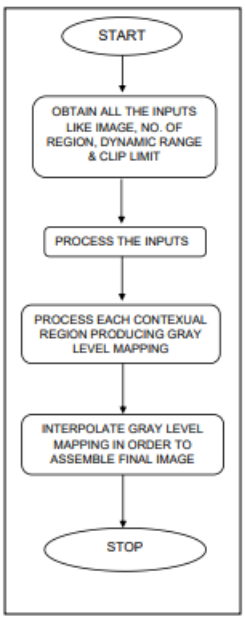

Fig.4 CLAHE Algo.

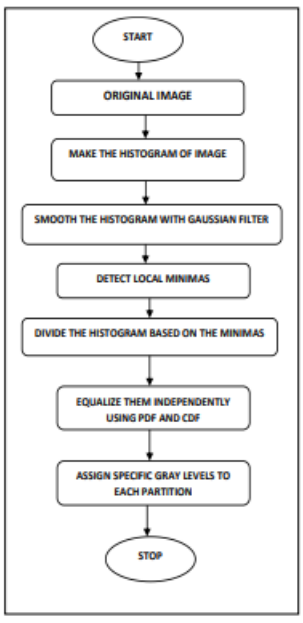

Fig.5 DHE Algo

Result and Discussions

A. Object Detection Results for YOLO V3 Framework:

The following table and chart show the detection results for YOLO V3 framework on COCO dataset on several images performed using trained model on COCO dataset. The results show detection results on several images with several objects detected with greater accuracy and inference time.

Sr.no.	Image	Inference Time (sec)	Detection Rate (%)			
			Object1	Object2	Object3	Object4
1	dog.jpg	00.91	99	99	94	
2	P_20170710_194702.jpg	01.41	99	99	94	65
3	P_20170825_214908.jpg	01.29	99	-	-	-
4	P_20170825_214912.jpg	01.28	99	-	-	-
5	Person.jpg	00.91	99	99	-	-
6	Room.jpg	00.90	99	99	99	-
7	Street.jpg	00.90	99	99	97	95

Fig.6 Table for Object Detection Values

B. Image Enhancement results using several color based enhancement techniques:

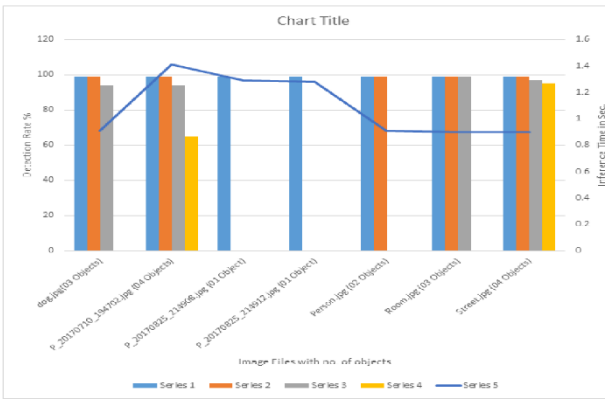

Fig.7 Object Detection results for YOLO V3 on several images.

The image enhancement performed on sample images using several image enhancement techniques like Contrast exposure fusion algorithm, Dynamic histogram equalization (DHE) algorithm and Contrast limited adaptive histogram equalization technique (CLAHE) are shown below, the results are obtained using histogram analysis method determining the color variation and pixel distribution of sample image and output enhanced image.

The results show that these techniques can be efficiently used in many low light and poor visual data conditions and are useful to improve the quality of visual data up to marginal extent. The histogram chart in which the x axis represent the pixel values, whereas y axis represent the color variance from dark to light.

i) Contrast Exposure Fusion Algorithm:

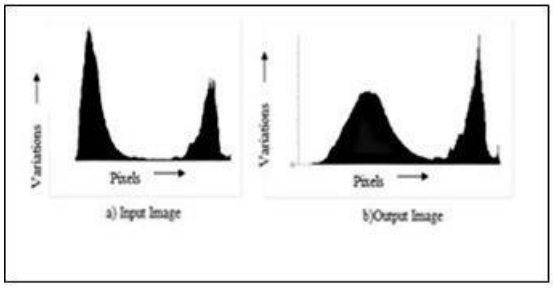

Fig.8 Histogram Results for sample Image CEFF Algo.

.

iii) Contrast Limited Adaptive Histogram Equalization (CLAHE):

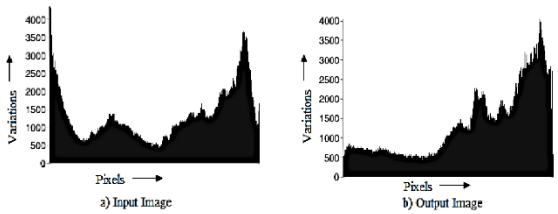

Fig.9 Histogram Results for sample Image DHE Algo

iii) Contrast Limited Adaptive Histogram Equalization (CLAHE):

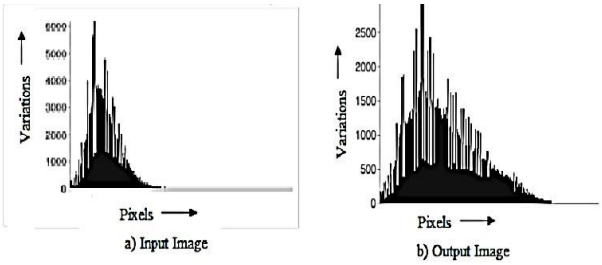

Fig.10 Histogram Results for sample Image CLAHE Algo.

Conclusion

Here we propose to develop a forensic video/Image analysis framework in order to perform truthful forensic analysis of visual data. We have implemented few modules such as of object detection using deep learning Darknet YOLO3 framework on COCO dataset, various contrast and histogram equalization algorithms for the purpose of image enhancement especially in low quality / low light type of image enhancement. The CNN based head pose estimation module can be useful in detection and identification of suspects in video footages. The other modules of framework such as skin detection using HSV range color detector, color detection using backprojection algorithm can be useful in color based analysis of visual data using color as finger-print, which are termed to be effective pattern recognition techniques. Also the techniques of inter-intra frame tampering detection can be used to determine that the available visual data is verifiably true. The above mentioned techniques can be used to perform visual data analysis in different scenarios of forensic analysis of visual data, surveillance areas, deep learning modules can even be used in robotics. The results until show that Yolo V3 detection model is termed to be the best detection model in single stage detectors, whereas the techniques like Contrast Limited Adaptive Histogram Equalization and Contrast Exposure Fusion Framework outperforms the other image enhancement techniques. In the future work, the different techniques and architecture of deep learning can be combined to gain more accurate and conclusive results. The future artificial intelligent (AI) systems in this specified area could reduce the gap between the human perception of visual data processing and artificial systems of visual data processing.

Acknowledgements

I dedicate my work to esteemed guide, Dr. S.K.Korde, whose interest and guidance helped me to complete the paper successfully. This experience will always steer me to my work perfectly and professionally. I would like to thank Dr. S. B. Tambe (PG Co-ordinator) who has provided facilities to explore the subject with more enthusiasm. I express my immense pleasure and thankfulness to all the teachers and staff of the Department of Comp. Engg., P.R.E.C, LONI for their cooperation and support. Last but not the least, I thank all others, and especially my friends who in one way or another helped me in the successful completion of this paper.

References

[1] Video-Based Evidence Analysis and Extraction in Digital Forensic Investigation, Jianyu Xiao, Shancang Li , And Qingliang Xu., 2019

[2] Head pose estimation in the wild using Convolutional Neural Networks and adaptive gradient methods, Massimiliano Patacchiola, Angelo Cangelosi, 2017.

[3] Yakun Chang, Cheolkon Jung, (Member, Ieee), Peng Ke, Hyoseob Song, And Jungmee Hwang, "Automatic Contrast-Limited Adaptive Histogram Equalization With Dual Gamma Correction". [4]"Head Pose Estimation in the Wild Assisted by Facial Landmarks Based on Convolutional Neural Networks", jiahao xia, libo cao, guanjun zhang, and jiacai liao – ieee 2019.

[5] S. Park, S. Yu, M. Kim, K. Park, and J. Paik, ``Dual autoencoder network for retinex-based low-light image enhancement,'' IEEE Access, vol. 6, pp. 22084-22093, 2018.

[6] W. Fan, K.Wang, C. François, and Z. Xiong, ``Median filtered image quality enhancement and anti-forensics via variational deconvolution,'' IEEE Trans. Inf. Forensics Security, vol. 10, no. 5, pp. 1076-1091, May 2015.

[7] C.-Y. Li, J.-C. Guo, R.-M. Cong, Y.-W. Pang, and B. Wang, ``Underwater image enhancement by dehazing with minimum information loss and histogram distribution prior,'' IEEE Trans. Image Process., vol. 25, no. 12, pp. 5664-5677, Dec. 2016.

[8] S. Mandal, X. L. Deán-Ben, and D. Razansky, ``Visual quality enhancement in optoacoustic tomography using active contour segmentation priors,'' IEEE Trans. Med. Imag., vol. 35, no. 10, pp. 2209_2217, Oct. 2016.

[9] S. Kim, W. Kang, E. Lee, and J. Paik, ``Wavelet-domain color image enhancement using ltered directional bases and frequencyadaptive shrinkage,'' IEEE Trans. Consum. Electron. vol. 56, no. 2, pp. 063-1070, May 2010.
[10] Kseniia Nikolskaia, Nadezhda Ezhova, Anton Sinkov, and Maksim, ``Skin Detection Technique Based on HSV Color Model and SLIC Segmentation Method,'' CEUR-WS.org/vol-2281/paper-13.pdf. [11] Y. Chang, C. Jung, P. Ke, H. Song, and J. Hwang, ``Automatic contrast limited adaptive histogram equalization with dual gamma correction,'' IEEE Access, vol. 6, pp. 11782-11792, 2018. [12] M. Grega, A. Matiola«ski, P. Guzik, and M. Leszczuk, ``Automated detection of firearms and knives in a CCTV image,'' Sensors, vol. 16, no. 1, p. 47, 2016.
[13] Graupe, Daniel, "Principle of artificial Neural networks", 2013, World Scientific Publishing Co Pte Ltd
[14] I. Goodfellow, Y. Bengio and A. Courville, Deep Learning, 2016, The MIT Press, Cambridge, MA, USA.
[15] Kim, Jinho. "Degraded Number Plate Image Recognition for CCTV Surveillance."
[16] "Digital Image Processing", R. C. Gonzalez & R. E. Woods, Addison-Wesley Publishing Company, Inc., 1992.
[17] Sitara, K., and B. M. Mehtre. "Digital video tampering detection: An overview of passive techniques." in Digital Investigation 18 (2016): 8-22.
[18] Wang, Wan, et al. "Identifying video forgery process using optical flow." International Workshop on Digital Watermarking. Springer, Berlin, Heidelberg, 2013.
[19] Chao, Juan, Xinghao Jiang, and Tanfeng Sun. "A novel video interframe forgery model detection scheme based on optical flow consistency." The International Workshop on Digital Forensics and Watermarking 2012. Springer, Berlin, Heidelberg, 2013.
[20] [20] Wang, Qi, et al. "Video inter-frame forgery identification based on optical flow consistency." Sensors & Transducers 166.3 (2014): 229.
[21] Wu, Yuxing, et al. "Exposing video inter-frame forgery based on velocity field consistency." Acoustics, speech and signal processing (ICASSP), 2014 IEEE International Conference on. IEEE, 2014.
[21] [22] Wang, Weihong, and Hany Farid. "Exposing digital forgeries in video by detecting double MPEG compression." Proceedings of the 8th workshop on Multimedia and security. ACM, 2006. 66
[23] Fridrich, A. Jessica, B. David Soukal, and A. Jan Lukáš. "Detection of copy-move forgery in digital images." in Proceedings of Digital Forensic Research Workshop. 2003.
[24] Amerini, Irene, et al. "A sift-based forensic method for copy–move attack detection and transformation recovery." IEEE Transactions on Information Forensics and Security 6.3 (2011): 1099-1110.
[25] Zampoglou, Markos, Symeon Papadopoulos, and Yiannis
Kompatsiaris. "Large-scale evaluation of splicing localization algorithms for web images." Multimedia Tools and Applications76.4 (2017): 48014834.
[26] Lin, Zhouchen, et al. "Fast, automatic and fine-grained tampered JPEG image detection via DCT coefficient analysis." Pattern Recognition 42.11 (2009): 2492-2501.
[27] Ferrara P, Bianchi T, De Rosa A, Piva A "Image forgery localization via fine-grained analysis of CFA artifacts." IEEE Transactions on Information Forensics and Security (2012) 7:1566–1577 [28] Mahdian B, Saic S "Using noise inconsistencies for blind image forensics." Image Vis Comput (2009) 27:1497–1503.
[29] J. Redmon, " YOLOv3: An Incremental Improvement". Available online https://pjreddie.com/yolo/.
[30] Zhenqiang Ying, Ge Li, Yurui Ren, Ronggang Wang, Wenmin Wang "A New Image Contrast Enhancement Algorithm using Exposure Fusion Framework".