



## A MACHINE LEARNING BASED SYSTEM TO ADDRESS THE FORENSIC ANALYSIS OF SCANNER DEVICES

**Assoc. CH.V. MURALI KRISHNA<sup>1</sup>**

**G. JAHNAVI<sup>2</sup>, A.ANU DEEPIKA<sup>3</sup>, SK. SANA BEGUM<sup>4</sup>, P. RAMESH<sup>5</sup>**

<sup>1</sup>Professor, Department of CSE (DATA SCIENCE), NRI INSTITUTE OF TECHNOLOGY, Vijayawada, A.P., India.

<sup>2, 3, 4, 5</sup>Student, Department of CSE (DATA SCIENCE), NRI INSTITUTE OF TECHNOLOGY, Vijayawada, A.P., India.

**Abstract** — a number of limitations and vulnerabilities associated with machine learning-based approaches, for example how to detect adversarial (image) examples, with real-world consequences (e.g., inadmissible evidence, or wrongful conviction). Therefore, with a focus on image forensics, this paper surveys techniques that can be used to enhance the robustness of machine learning-based binary manipulation detectors in various adversarial scenarios. many forensic techniques such as digital image authentication, source identification and tamper detection are important for forensic image analysis.

### INTRODUCTION

In this project, we describe a machine learning based system to address the forensic analysis of scanner devices. The proposed

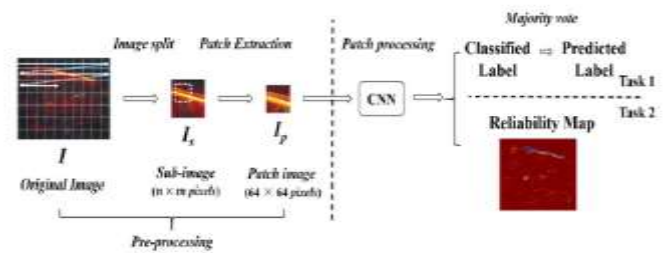
system uses machine-learning to automatically learn the intrinsic features from various scanned images. These scanner independent features increase the difficulty in scanner model identification. Many scanners also use 1D “line” sensors, which are different than the 2D “area” sensors used in cameras. Previous work in scanner classification and scanned image forensics mainly focus on handcrafted feature extraction [9], [10], [11]. They extract features unrelated to image content, such as sensor pattern noise [9], dust and scratches [10]. In [12], Gou et al. extract statistical features from images and use principle component analysis (PCA) and support vector machine (SVM) to do scanner model identification. The goal is to classify an image based on scanner model rather than the exact instance of the image. In [9], linear discriminant analysis (LDA) and SVM are

used with the features which describe the noise pattern of a scanned image to identify the scanner model. This method achieves high classification accuracy and is robust under various post-processing (e.g., contrast stretching and sharpening). In [10], Dirik et al. propose to use the impurities (I.e, dirt) on the scanner pane to identify the scanning device.

### PROPOSED SYSTEM

Convolutional neural networks (CNNs) such as VGG [13], Res Net [14], Google Net [15], and Xception [16] have produced state-of-art results in object classification on ImageNet [17]. CNNs have large learning capacities to “describe” imaging sensor characteristics by capturing low/median/high-level features of images [8]. For this reason, they have been used for camera model identification [8], [18] and have achieved state-of-art results.

In this paper, we propose a CNN-based system for scanner model identification. We will investigate the reduction of the network depth and number of parameters to account for small image patches (i.e.,  $64 \times 64$  pixels) while keeping the time for training in a reasonable range. Inspired by [16], we propose a network that is light-weight and also combines the advantages of Res Net [14] and Google Net [15]. The proposed system can achieve a good classification accuracy and generate a reliability map (i.e., a heat map, to indicate the suspected manipulated region).



### EXISTING SYSTEM

With powerful image editing tools such as Photoshop and GIMP being easily accessible, image manipulation has become very easy. Hence, developing forensic tools to determine the origin or verify the authenticity of a digital image is important. These tools provide an indication as to whether an image is modified and the region where the modification has occurred. A number of methods have been developed for digital image forensics. For example, forensic tools have been developed to detect copy-move attacks and splicing attacks.

### DISADVANTAGES

Less accuracy.

### PROPOSED SYSTEM

The proposed system an input image is first split into smaller sub-images  $I_s$  of size  $n \times m$  pixels. This is done for four reasons: a) to deal with large scanned images at native resolution, b) to take location independence into account, c) to enlarge the dataset, and d) to provide low pre-processing time.

### ADVANTAGES

More accurate

### LITERATURE SURVEY



### **Camera model identification with the use of deep convolutional neural networks**

In this paper, we propose a camera model identification method based on deep convolutional neural networks (CNNs). Unlike traditional methods, CNNs can automatically and simultaneously extract features and learn to classify during the learning process. A layer of preprocessing is added to the CNN model, and consists of a high pass filter which is applied to the input image. Before feeding the CNN, we examined the CNN model with two types of residuals. The convolution and classification are then processed inside the network. The CNN outputs an identification score for each camera model. Experimental comparison with a classical two steps machine learning approach shows that the proposed method can achieve significant detection performance. The well known object recognition CNN models, Alex Net and Google Net, are also examined.

### **Digital camera identification from sensor pattern noise**

In this paper, we propose a new method for the problem of digital camera identification from its images based on the sensor's pattern noise. For each camera under investigation, we first determine its reference pattern noise, which serves as a unique identification fingerprint. This is achieved by averaging the noise obtained from multiple images using a denoising filter. To identify the camera from a given image, we consider the reference pattern noise as a spread-spectrum watermark,

whose presence in the image is established by using a correlation detector. Experiments on approximately 320 images taken with nine alarm rates and false rejection rates. Additionally, we study how the error rates change with common image processing, such as JPEG compression or gamma correction.

**M. Kharrazi et al. (2004).** The interpolation in the colour surface of an image due to the use of a colour filter array (CFA) forms the basis of the paper. We propose to identify the source camera of an image based on traces of the proprietary interpolation algorithm deployed by a digital camera. For this purpose, a set of image characteristics are defined and then used in conjunction with a support vector machine based multi-class classifier to determine the originating digital camera. We also provide initial results on identifying source among two and three digital cameras.

### **RELATED WORK**

#### **TRAINING**

A test image will first be split into sub-images, and then subsequently extracted into patches of size 64×64 pixels. The extracted patches will be used as inputs for the proposed neural network.

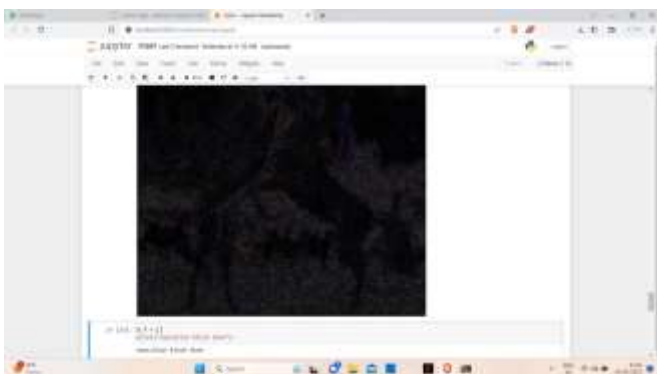
#### **PRE-PROCESSING**

This pre-processing enables the proposed system to work with small-size images and use smaller network architecture to save training time and memory usage. Designing suitable network architecture is an important part in the scanner model identification system.

## TESTING

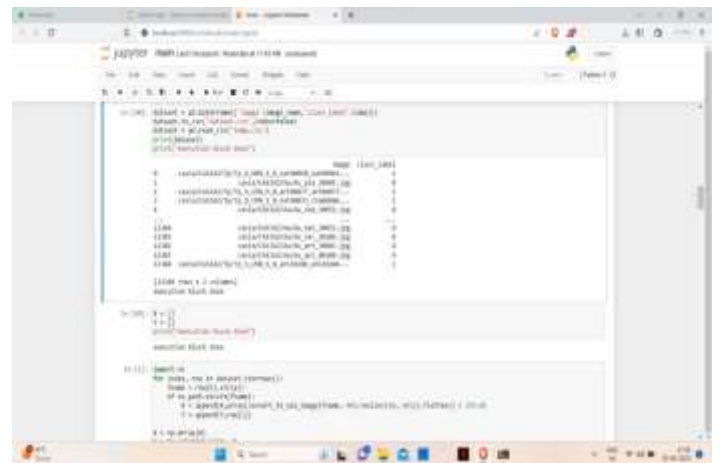
The same pre-processing procedure as described in the training section will be used in the testing stage. Our proposed system will evaluate two tasks on scanned images: scanner model classification and reliability map generation. In Task 1 (scanner model classification), we assign the predicted scanner labels to both patches  $I_p$  and original images  $I$ . The predicted scanner label for the sub-image is the same as the predicted label of its corresponding patch. The classification decision for the original image  $I$  is obtained by majority voting over the decisions corresponding to its individual sub-images. In Task 2, a reliability map [19] is generated based on the majority vote result from Task 1. The pixel values in the reliability map indicate the probability of the corresponding pixel in the original image being correctly classified.

## SAMPLE SCREEN SHOTS



## CONCLUSION

In this paper, maintain part of their effectiveness even against CNN models other than the one targeted by the attack. This is a very strong property undermining the usability of CNN 's security-oriented applications. In this paper, we investigate if attack transferability also holds in image forensics applications. With specific reference to the case of manipulation detection, we analyse the results of several experiments considering different sources of mismatch between the CNN.



## REFERENCES

- [1] Acquiring the user's opinion by using a generalized context-aware recommender system



for real- world applications CVM Krishna, DGA

Rao Int J Eng Technol 7 (2.7), 883-886

[2] Analysing the impact of contextual segments on the overall rating in multi-criteria

recommender systems CVM Krishna, GA Rao,

S Anuradha Journal of big Data 10 (1), 16

[3] Impact of Contextual Segments in the Prediction of Overall User Gratification in Asian and European Continental Hotel Tourism Sector

CVM Krishna, G Appa Rao, BB Kadaru,

S Anu Radha ICCCE 2021: Proceedings of the 4th International Conference on ...

[4] Prediction of User Overall Gratification in Indian Tourism Domain on Hotel Classes and Trip Types CVM Krishna, GA Rao, S Anu Radha,

KVD Sagar ECS Transactions 107 (1), 19813

[5] Image anonymization using deep convolutional generative adversarial network KN Rao, P

Jayasree, CVM Krishna, KS Prasanth, CS Reddy

Journal of Physics: Conference Series 2089 (1), 012012

[6] Predictive analytics of heart disease presence with feature importance based on machine

learning algorithms SD Nitalaksheswara Rao Kolukula<sup>1</sup>, Prathap Nayudu Pothineni<sup>2</sup>, Venkata ...

Indonesian Journal of Electrical Engineering and Computer Science 32 (2 ...)

[7] A framework for the identification of significant contexts in tourism domain CV Krishna, GA Rao, S.

Anu Radha Int J Adv Sci Technol 29 (7), 1007-29

[8] A study on prediction of user overall gratification in European continental tourism city

hotels CVM Krishna, M Srinath, KN Sri, KH

Kumar, GK Kumar AIP Conference Proceedings 2796 (1)

[9] Users preference assessment on overall

gratification in Asian tourism hotels before and

after pandemic CVM Krishna, KB Sri, S Avinash,

DA Priya, G Yamini AIP Conference Proceedings 2796 (1)

[10] Impact of user reviews and ratings on overall rating in Indian tourism hotels CVM Krishna, AC

Sri, S Ayesha, GR Lakshmi, DP Shankar

AIP Conference Proceedings 2796 (1)

[11] Impact of Contextual Segments in the

Prediction of Overall User Gratification in Asian and European Continental Hotel Tourism Sector

CVM Krishna, G Appa Rao, BB Kadaru,

S Anu Radha ICCCE 2021: Proceedings of the 4th International Conference on ...



## BIOGRAPHIES

**Ch. V. Murli Krishna** is an Associate Professor and the head of the CSE (Data Science) department at NRI Institute of Technology. He has over 20 years of experience in engineering academics and has taught both undergraduate and postgraduate courses. He has also guided many B. Tech, M. Tech, and MCA projects as part of the academic curriculum. In addition to his teaching responsibilities, he has been involved in various administrative tasks like NBA, NAAC, ICT, and IQAC. He has organized several workshops, faculty development programs, and Tech Fests. He is a member of IAENG, IFERP, and INSC.



He completed his M. Tech. in Computer Science & Engineering in 2009 and is currently pursuing a Ph.D. in Computer Science & Engineering at GITAM (Deemed to be University), Vizag.

**Atchuta Anu Deepika** is currently pursuing BTech with specification of Computer Science with Data Science Engineering in NRI Institute of Technology. She had completed her mini project on customer churn. She completed an internship with Blackbucks company and she had NPTEL certification on the Cloud computing.



**Shaik Sana Begum** is currently pursuing BTech with specification of Computer Science with Data Science Engineering in NRI Institute of Technology. She had completed her mini project on Migration Prediction. She completed an internship with Blackbucks company and she had NPTEL certification on the Cloud computing.





**Guntupalli Jahnavi** is currently pursuing BTech with specification of Computer Science with Data Science Engineering in NRI Institute of Technology. She had completed her mini project on Contact Tracing. She completed an internship with Blackbucks company and she had NPTEL certification on the Cloud Computing and Data Analytics with python.



**Pinniboina Ramesh** is currently pursuing BTech with specification of Computer Science with Data Science Engineering in NRI Institute of Technology. He had completed his mini project on Credit Card fraud detection. He completed an internship with Blackbucks company.

