Deep-RSI: Deep learning for Radiographs Source Identification

Farid Ghareh Mohammadi[1,2], Ronnie Sebro [1,2]

1 : Department of Radiology, Mayo Clinic, Jacksonville, Fl, 32224
2 :Center for Augmented Intelligence, Mayo Clinic, Jacksonville, Fl, 32224
Emails : Gharehmohammadi.farid@mayo.edu, Sebro.ronnie@mayo.edu

Abstract— *In forensics, the authenticity of digital images is of the utmost importance, considering that modern technology makes it incredibly simple and quick to alter and generate fake but convincing images. As a result, digital image credibility has decreased, making it difficult to demonstrate the source of images. Prior studies have shown that magnetic resonance imaging (MRI) scans can be traced back to their sources, but radiographs have not. In this paper, we propose the Deep-RSI algorithm, an algorithm that identifies the source (manufacturer and model) of the device used to create radiographs. This is the first time that a medical forensics investigation of this kind will be accomplished to declare and confirm radiograph sources. Researchers in information forensics, security, and medical imaging can use this data to determine scientific fraud, like fake radiographs made from unreliable sources or cut-and-paste fakes. This proposed solution describes how non-content pixels in images enable us to discover the manufacturer and model of a radiographic machine. Since radiographs are obtained from different sites of the body, source recognition has to be sensitive and free of any content-specific information. This will prevent the convolutional neural networks (CNN) from detecting content-specific details and instead identify fingerprints that are unique to the source. CNNs start with low-level features and, in the convolutional blocks, generate high-level features to identify the radiographic machine sources. This proposed solution reports the source (manufacturer and model) of each image. We obtain the highest AUC of 0.97 and a prediction accuracy of 98.54% for radiographic machine manufacturer detection. Our results show that forensic assessments of radiographs can be done with a high level of certainty.*

Keywords: *Deep Learning, Machine Learning, Radiographs Imaging, Medical Imaging, Forensic Imaging*

1. Introduction

Due to changes in technology, it is now easier than ever to change digital images. The forensics community has increasingly concerned with validating the authenticity of digital medical images [1], [2]. Inference based on the contents of images that are not an accurate representation of the patient may result in inaccurate clinical decisions [3]. There is little known about the difficulty of identifying the origin of digital medical images. Previous research on magnetic resonance imaging (MRI) scan source detection has shown promising results for detecting MRI sources [4], but no research has been conducted on identifying the source of radiographic machine sources. This novel research will provide a universal forensic method for accurately predicting the source (manufacturer and model) of radiographs using deep learning techniques. To our knowledge, this is a pioneering study in medical forensic research to identify and validate the source of radiographs. Researchers in information forensics, security, and medical imaging can utilize this data to uncover instances of scientific fraud, such as falsified radiographs that do not originate from a reliable source or that contain obvious cut-and-paste forgeries [5]. In addition, our findings contribute to the growing body of evidence demonstrating that forensic assessments may be conducted with high degrees of certainty.

Because fabricated images might deceive people, it is critical to research methods to verify whether a digital image is real or not [6]. Tracing the origins of radiographs may have a variety of significant benefits for both medical forensics and information forensics [7]. Researchers may be unable to protect patients' confidentiality and privacy by tracing the source of leaked radiographs. Researchers may also be able to secure medical information from unauthorized tampering by validating the provenance of certain radiographs.

Deep learning techniques use convolutional neural networks (CNNs), which have more than two hidden layers and extra filters. Their design is based on stacking several hidden layers, and consists of the characteristics and components required to establish a network [8]. This has been shown to be a good way to get hierarchical information from the foregrounds and backgrounds of images. In other words, they can learn new things from a group of things they already know [9]. The first layer of a CNN architecture is made up of a set of convolutional feature extractors that are applied to the image in parallel by a set of multiple learnable filters. To make feature maps, these filters work like a sliding window that interacts with all parts of the image by a certain amount, called the stride. In the same way, the hidden convolutional layers pull out each lower-level feature map. The output of these hierarchical feature extractors is then fed into a fully connected neural network that sorts things into classification.

979-8-3503-2028-2/22/$31.00 ©2022 IEEE
DOI 10.1109/CSCI58124.2022.00307

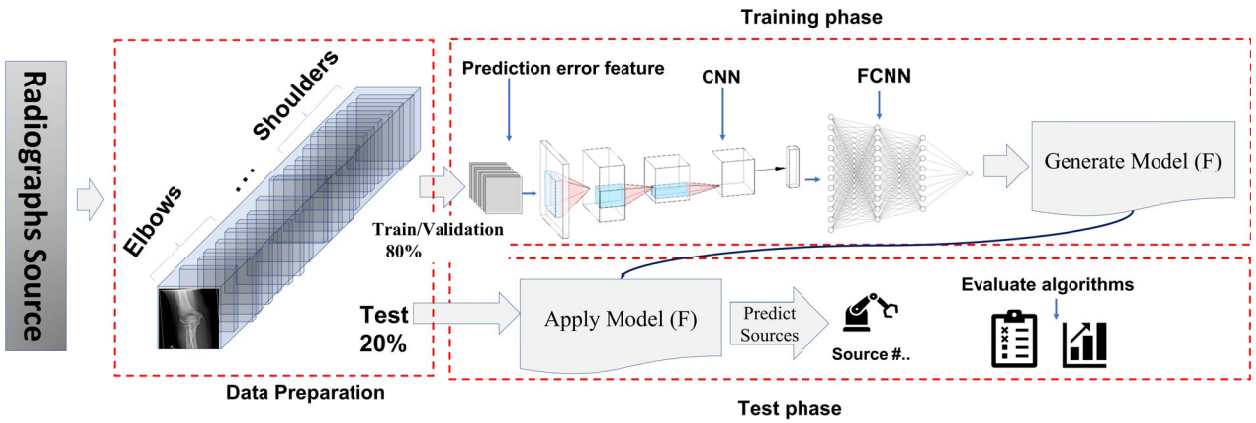

Fig. 1

A GENERAL SCHEMA OF RADIOGRAPHS SOURCE IDENTIFICATION USING DEEP-RSI

However, CNNs with their current networks and settings didn't recognize the altered or computer-made images. In this research, we present a new approach for addressing this problem by utilizing new CNN configurations.

In this paper, we propose a novel technique, deep learning for radiographs source identification (Deep-RSI), in which we use a customized convolutional neural network (CNN) to learn the data from non-content pixels and predict radiograph manufacturers and models. The operation uses CNN layers and filters to extract low-level (pixel-level) data from noise and content-free images located in the background, and then CNN learns high-level features. This should allow CNN to skip anatomical aspects (image content) inside an image and concentrate entirely on radiographics' machine traces. Then, fully connected layers of convolutional neural networks (FCNN) learn the data generated by the combination of the first layer followed by convolutional blocks to get high-level features. This automated machine learning algorithm provides a manufacturer and model identification for each image.

Recent advancements in digital imaging technologies have given rise to a new set of issues and concerns addressing image authenticity [10]. It is now possible to create digital images, edit them, and manipulate them without leaving any visible evidence of the actions taken. These characteristics substantially undermine the trustworthiness of images. A new field of study known as "digital image forensics" aims to determine the origin and likely authenticity of a digital image. The main contribution of this proposed research is to learn how the content-free pixels in images can provide information about their origins, such as the radiographic machine's manufacturer and model. It is crucial to achieve source recognition through a sensitive, content-free operation since diverse parts of the body are scanned to create radiographs. The major contributions of our research are as follows:

◇ We use radiographs to identify radiographic machine source with high accuracy.

◇ We use radiographs to identify radiographic machine source with high values for area under curve (AUC) for each model.

◇ We leverage content-free pixels to extract low-level features and identify radiographic machine sources fingerprints.

◇ We detect and extract fingerprints automatically out of radiographs for the source identification.

◇ We customize sampling data during the training phase to learn an unbalanced dataset fairly using a new approach for data weighting per class (radiographic machine model / manufacturer) to avoid having a biased model for the majority of classes.

The rest of this paper is organized as follows: In Section 2, we provide an overview of recent forensic studies to distinguish altered images. Then, in Section 3, we elaborate on Deep-RSI architecture and discuss how CNNs are applied in medical forensics tasks. Finally, we evaluate the Deep-RSI results and analyze them.

2. Related Work

Radiographic machine with different camera models and setups make output images with different resolutions, which shows a relationship between non-content pixels in the images [12], [13]. Pattern noise (fingerprint) is introduced into digital images by various digital cameras. The cause of pattern noise is the imperfection of embedded sensors, their device equipment, and in-camera preparatory activities, resulting in a range of fingerprint patterns in the images [14]. As a result, scientists can use these fingerprints to distinguish images (falsified images from authentic images) based on their sources.

Table 1
DEEP-RSI THEME FOR SOURCE IDENTIFICATION BUILD UPON MISLNET [11]

Layer	Info
Conv1	Conv2d(3, 96, kernel size=7, stride=2, pad.=4)
Max pool	MaxPool2d(kernel size=3, stride=2, pad.=0)
Conv2	Conv2d(96, 64, kernel size=5, stride=1, pad.=2)
Max pool	MaxPool2d(kernel size=3, stride=2, pad.=0)
Conv3	Conv2d(64, 64, kernel size=5, stride=1, pad.=2)
Max pool	MaxPool2d(kernel size=3, stride=2, ,pad.=0)
Conv4	Conv2d(64, 128, kernel size=1, stride=1,pad.=0)
Avg pool	AvgPool2d(kernel size=3, stride=2, pad.=0)
FCNN1	Linear(in features=6272, out features=200)
FCNN2	Linear(in features=200, out features=200)
FCNN3	Linear(in features=200, out features=classes)

Experts in digital and multimedia forensic science have developed a variety of ways for detecting the manufacturer and model of digital photos [15] and video sources [16], ranging from more traditional approaches to ones that depend on deep learning. Because malware assaults on healthcare systems are becoming more widespread, similar approaches have been extended to identify the sources of magnetic resonance imaging (MRI) scans [4].

In image forensics, the difficulty of distinguishing fabricated images from authentic images can be divided into two categories: hand-crafted feature-based methods and deep learning-based methods. The former focuses primarily on the physical differentiation between faked and actual images throughout the image synthesis process [17], and is further separated into statistical data-based identification approaches [18] and physical features [19]. Furthermore, Villalba *et al* proposed a method for video source identification based on fingerprint extraction from video key frames [20]. Manually created features, on the other hand, generally result in huge feature dimensions, complicated calculations, and low detection rates [3]. To address this problem, researchers proposed potential solutions by employing feature selection methods [6], [21], which boosted the detection rate and computation time while considerably decreasing the feature dimension.

The later category, deep learning algorithms, generates features automatically using CNNs from images to feed classifiers that learn and predict the images' source. In recent years, deep learning-based algorithms have made substantial progress [4]. The CNN-based techniques primarily use a classifier to extract hierarchical representations from the source image and differentiate between falsified images and real images [4], [22], [23].

3. Proposed Method

There is a high degree of association between non-content pixels that are adjacent to one another. A decent steganalysis will look for a harmony that may be found between the pixels that are adjacent to one another [21]. Deep-RSI seeks to learn fingerprints to detect radiographs' source (manufacturers and models of the radiographic machine). To that end, we elaborate on three main sections that enable us to fulfill the goal of Deep-RSI.

3.1 Data Preparation

We extract all digital radiographs captured in Mayo Clinic Florida (MCF) from 2010-2021 for five sites: hands, wrists, forearms, elbows, and shoulders to perform this research study. In addition, we include every manufacturer and model list that has been active on the MCF campus since 2010. The radiographs format is Digital Imaging and Communications in Medicine (DICOM), which provides information about patients, device stations, specialists, nurses, and operators, as well as the configurations of the output images and the location of the room where the radiographs were captured, as well as an image of the patient's site itself. We extract the manufacturer and model information from DICOMs and compare it to the list.

Next, we apply the pre-processing technique to remove ineligible DICOMs, which are small in size and lacking in information, from the dataset with some conditions. We removing unknown DICOMs which are related to device quality control (QC). We report 7898 DICOMs with 10 classes of radiographic machine model and four classes of radiographic machine manufacturers. We call model0 up to model9 and manufacturer0 up to manufacturer3 due to Health Insurance Portability and Accountability Act (HIPAA) policy. Finally, we convert all DICOMs to Joint Photographic Experts Group (JPEGs /JPGs) format and resize the JPGs into 256*256 in order to fit the proposed architecture and model. Lastly, we randomly split the dataset into training and test datasets (80/20) so that the number of manufacturers and models were balanced in the training and test datasets. We allocate 6318 sample JPGs for training and 1580 for the test dataset.

3.2 Data Weighting

The information in table 2 demonstrates that we have an unbalanced dataset to learn from in order to detect radiographic machine models and manufacturers. The standard sample splitting procedure for training and testing yields an unbalanced amount of samples for each class (radiographic machine model / manufacturer). We shuffle the samples in the dataset, but this does not yield a promising result for the learning model using the unbalanced dataset. In this research, we present a novel method for leveraging the weight of each class in comparison to the entire dataset. Then, while training, we select dataset samples based on the weights of the classes (radiographic machine model / manufacturer). This allows us to select samples fairly for training and testing, ensuring that the learned model is not biased toward classes with the majority of examples.

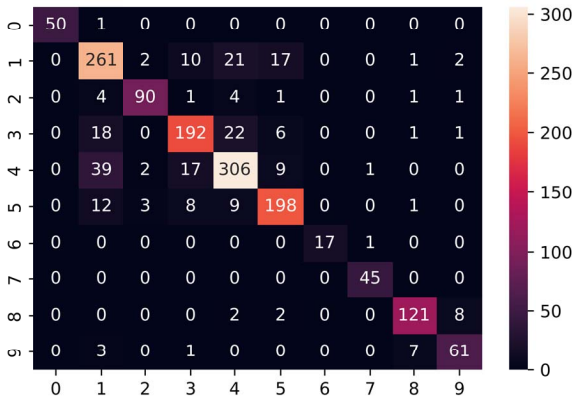

Fig. 2

A CONFUSION MATRIX FOR THE MODELS IDENTIFICATION

3.3 Training Phase

We present the Deep-RSI Network architecture in figure 1 in which you see the training phase that takes the training dataset to learn and make a model (F). We were inspired by the work [4] in this customized network, and built Deep-RSI on the CNN architecture proposed by Bayar and Stamm in "Bayar 2018 Constraints." We first generated a "Residual" or "Prediction Error Feature" layer that provided all content-free pixel information. To generate this layer, we use a filter 5*5, with the information appearing in the formula below:

$$W_c^1 = \begin{cases} w_c^1(x,y) = -1 & x,y = 0 \\ \sum w_c^1(x,y) = 1 & -3 < x,y < 3, \neq 0 \end{cases} \quad (1)$$

where $w_c^1(x,y)$ is the filter weight at the (x, y) position and $w_c^1(0,0)$ is the filter weight at the center of the filter window (W). Each filter in *Residual* layer is initialized by randomly chosen, then enforcing the limitations in equation 1. During training, the restrictions in equation 1 are applied again after the filter weights have been updated using stochastic gradient descent. This equation enables the CNN to train a robust collection of alteration detection feature extractors adaptively, rather than having them pre-selected. Because these features are low-level and easily round up/down, causing us to lose information about non-content pixels, we do not apply any pooling for this layer after convolution has completed.

Following this layer are three convolutional blocks, each of which has numerous convolutional filters, batch normalization, tanh activation, and max pooling. These convolutional blocks are intended to extract high-level information from low-level noisy input produced by the restricted layer. A learned forensics feature extractor is created by combining a restricted layer with many convolutional blocks. Following the convolutional blocks, there are two fully connected neural network (FCNN) layers with 200 neurons each, followed by an output layer with softmax activation, where each neuron corresponds to a single class (radiographic machine model/manufacturer). We use the FCNN as our main classifier to predict ten radiographic machine models (10 classes) and four radiographic machine manufacturers (4 classes).

We choose this architecture [11] as our baseline because its compressed convolutional laers are designed to suppress visual content and learn from forensic evidence. This allows the CNN to disregard anatomical details in a scan and concentrate entirely on radiographs fingerprints.

We set Deep-RSI's hyperparameters as follows. The number of classes is 10 (radiographic machine models) and 4 (radiographic machine manufacturers), the number of epochs is 100, and the learning rate is 0.0001. Table shows the remaining configurations.1.

We train CNNs to accomplish two main goals. First, radiographs machine manufacturer identification, and second, radiograph model identification leveraging Deep-RSI.

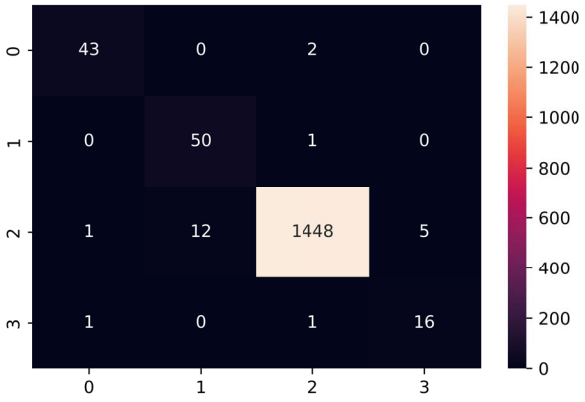

Fig. 3

A CONFUSION MATRIX FOR THE MANUFACTURER IDENTIFICATION

3.4 Test Phase

After completing the training phase and generating the model (F), we evaluate F on the unbiased test dataset. This is why the test and training datasets are separated to avoid a biased model. In the following section, we report the Deep-RSI performance for radiograph manufacturers and model identification.

4. Experimental Results

In this section, we present our experimental setup as well as the results of our research on radiographic machine

Table 2
RADIOGRAPHIC MACHINE MODEL DISTRIBUTION

Models	Total	Training	Test
Model 0	215	164	51
Model 1	1588	1274	314
Model 2	447	345	102
Model 3	1363	1123	240
Model 4	1778	1404	374
Model 5	1180	949	231
Model 6	85	67	18
Model 7	200	155	45
Model 8	758	625	133
Model 9	284	212	72
Total	7898	6318	1580

manufacturer and model identification. We construct training and test datasets of radiograph scans and then analyze Deep-RSI's performance using the test dataset. Our results show that the radiographic machine does, in fact, leave traces that can be found and used by the forensic CNN model to determine the source of the radiographs.

4.1 Dataset

We collect all radiographs source manufacturers, models, and radiographs collected at Mayo Clinic Florida from 2010 for five sites: hands, wrists, forearms, elbows, and shoulders. The total number of samples is 7898, 6318 of which are in training, and 1580 of which are in testing. Table 2 illustrates that we have ten different radiographic machine models that create DICOMs for the various sites. We have a minimum of 85 for the model six and a maximum of 1588 for the model one.

4.2 Radiographic machine models evaluation

We begin by evaluating Deep-RSI on radiographic machine models using three key metrics: TPR (or sensitivity), TNR (or specificity), and AUC (area under the curve). We must compute the AUC for a multi-class dataset because the classes in this dataset are not binary. To accomplish this, we must compare each class to others and calculate AUC for each class. According to Table 3 our proposed method produces promising results for each radiographic machine's model detection. The maximum AUC is 0.99 for radiographic machine model zero, and the lowest AUC is 0.88 for radiographic machine models one and four. We achieve 100% sensitivity for radiographic machine model seven and 100% specificity for radiographic machine models zero and six. The total accuracy of radiographic machine model detection for all radiographic machine models is calculated to be 84.87 percent.

We generate a confusion matrix for radiographic machine model detection that is shown in figure 2. This figure shows that radiographic machine model#4 has the highest number of test samples for model number four in which we correctly

Table 3
STATISTICAL RESULTS FOR MODELS

Models	TPR(%)	TNR (%)	AUC
Model 0	98.04	100	0.99
Model 1	83.12	93.92	0.88
Model 2	88.23	99.53	0.94
Model 3	80.0	97.24	0.89
Model 4	81.81	95.19	0.88
Model 5	85.71	97.40	0.91
Model 6	94.44	100	0.97
Model 7	100	99.87	0.99
Model 8	90.97	99.24	0.95
Model 9	84.72	99.20	0.92

Table 4
RADIOGRAPHIC MACHINE MODELS IDENTIFICATION RESULTS

Sites	Total samples	Accuracy(%)
Hands	3949	85.95
Wrists	386	79.48
Forearms	61	84.61
Elbows	268	90.74
Shoulders	3265	86.37
Total	7898	84.87

detect 306 the samples, with AUC value of 0.88. note that we do not calculate the radiographic machine manufacturer detection for each site since we only have one radiographic machine manufacturer for each site for more than 99 percent of the samples.

We evaluate our proposed method on five separate sites for radiographic machine model detection. Table 4 shows the number of samples as well as the prediction accuracy for each site. The model's low accuracy is due to a lack of samples per class (radiographic machine model) during the training phase. Despite the fact that the hands-on class contains 3949 samples, fewer radiographs are captured. As a result of this, the trained model failed to recognize particular radiographic machine models for the site.

4.3 Radiographic machine manufacturer evaluation

Table 5 shows that we have four different radiographic machine manufacturers (0 to 3) in which the manufacturer # 2 has the majority of samples in comparison with others, and the manufacturer #3 has minimum number of samples. We have to make sure that the training process fairly consider all radiographic machine manufacturers (classes) to avoid biased information on the manufacturer #2. We leverage a data weighting approach to obtain a promising result of 98.54% accuracy in radiographic machine manufacturer detection.

In addition, we project the Deep-RSI results as a confusion matrix onto a heat-map chart, as shown in figure 3. This

Table 5
RADIOGRAPHIC MACHINE MANUFACTURER DISTRIBUTION

Models	Total	Training	Test	TPR(%)	TNR (%)	AUC
Manufacturer 0	200	155	45	95.56	99.87	0.98
Manufacturer 1	215	164	51	98.04	99.21	0.99
Manufacturer 2	7398	5932	1466	98.77	96.49	0.98
Manufacturer 3	85	67	18	88.89	99.68	0.94
Total	7898	6318	1580	98.54	96.74	0.98

graph shows that the manufacturer #2 represents the bulk of radiographic machine manufacturers.

5. Conclusion

In this paper, we demonstrate that it is possible to identify the origin of sources of radiographs (manufacturer and model of the radiographic machine) that produced them. There has been no research done on the topic of identifying the original sources of radiographic machines using radiographs, the first research of this type in medical imaging. Researchers in information forensics, security, and medical imaging may consider this finding paramount. We also report the capability of a CNN to determine the manufacturer and model of the radiographic machine that was used to create the radiographs. By applying our proposed method to a new set of radiographs collected at the Mayo Clinic in Florida, we claim that we can correctly determine the manufacturer of the radiographic machine used to create the images with a success rate of 98.54% and the specific model with an accuracy rate of 84.87%.

References

[1] F. G. Mohammadi and M. S. Abadeh, "A survey of data mining techniques for steganalysis," in *Recent advances in steganography*. IntechOpen, 2012, no. 1.

[2] D. Gong, O. S. Goh, Y. J. Kumar, Z. Ye, and W. Chi, "Deepfake forensics, an ai-synthesized detection with deep convolutional generative adversarial networks," *International Journal*, vol. 9, no. 3, 2020.

[3] S. Mandelli, N. Bonettini, and P. Bestagini, "Source camera model identification," in *Multimedia Forensics*. Springer, Singapore, 2022, pp. 133–173.

[4] S. Fang, R. A. Sebro, and M. C. Stamm, "A deep learning approach to mri scanner manufacturer and model identification," *Electronic Imaging*, vol. 2020, no. 4, pp. 217–1, 2020.

[5] A. Ghai, P. Kumar, and S. Gupta, "A deep-learning-based image forgery detection framework for controlling the spread of misinformation," *Information Technology & People*, 2021.

[6] F. G. Mohammadi and M. S. Abadeh, "Image steganalysis using a bee colony based feature selection algorithm," *Engineering Applications of Artificial Intelligence*, vol. 31, pp. 35–43, 2014.

[7] D. P. Chowdhury, S. Bakshi, P. K. Sa, and B. Majhi, "Wavelet energy feature based source camera identification for ear biometric images," *Pattern Recognition Letters*, vol. 130, pp. 139–147, 2020.

[8] F. Shenavarmasouleh and H. R. Arabnia, "Drdr: Automatic masking of exudates and microaneurysms caused by diabetic retinopathy using mask r-cnn and transfer learning," in *Advances in computer vision and computational biology*. Springer, 2021, pp. 307–318.

[9] F. Shenavarmasouleh, F. G. Mohammadi, M. H. Amini, T. Taha, K. Rasheed, and H. R. Arabnia, "Drdrv3: Complete lesion detection in fundus images using mask r-cnn, transfer learning, and lstm," *arXiv preprint arXiv:2108.08095*, 2021.

[10] M. C. Stamm, M. Wu, and K. R. Liu, "Information forensics: An overview of the first decade," *IEEE access*, vol. 1, pp. 167–200, 2013.

[11] B. Bayar and M. C. Stamm, "Constrained convolutional neural networks: A new approach towards general purpose image manipulation detection," *IEEE Transactions on Information Forensics and Security*, vol. 13, no. 11, pp. 2691–2706, 2018.

[12] X. Zhao and M. C. Stamm, "Computationally efficient demosaicing filter estimation for forensic camera model identification," in *2016 IEEE international conference on image processing (ICIP)*. IEEE, 2016, pp. 151–155.

[13] B. Bayar and M. C. Stamm, "A deep learning approach to universal image manipulation detection using a new convolutional layer," in *Proceedings of the 4th ACM workshop on information hiding and multimedia security*, 2016, pp. 5–10.

[14] G. Wu, X. Kang, and K. R. Liu, "A context adaptive predictor of sensor pattern noise for camera source identification," in *2012 19th IEEE International Conference on Image Processing*. IEEE, 2012, pp. 237–240.

[15] L. Bondi, L. Baroffio, D. Güera, P. Bestagini, E. J. Delp, and S. Tubaro, "First steps toward camera model identification with convolutional neural networks," *IEEE Signal Processing Letters*, vol. 24, no. 3, pp. 259–263, 2016.

[16] B. Hosler, O. Mayer, B. Bayar, X. Zhao, C. Chen, J. A. Shackleford, and M. C. Stamm, "A video camera model identification system using deep learning and fusion," in *ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2019, pp. 8271–8275.

[17] C.-H. Choi, H.-Y. Lee, and H.-K. Lee, "Estimation of color modification in digital images by cfa pattern change," *Forensic science international*, vol. 226, no. 1-3, pp. 94–105, 2013.

[18] F. Gisolf, P. Barens, E. Snel, A. Malgoezar, M. Vos, A. Mieremet, and Z. Geradts, "Common source identification of images in large databases," *Forensic science international*, vol. 244, pp. 222–230, 2014.

[19] Y. Yao, Z. Zhang, X. Ni, Z. Shen, L. Chen, and D. Xu, "Cgnet: Detecting computer-generated images based on transfer learning with attention module," *Signal Processing: Image Communication*, vol. 105, p. 116692, 2022.

[20] L. J. G. Villalba, A. L. S. Orozco, R. R. López, and J. H. Castro, "Identification of smartphone brand and model via forensic video analysis," *Expert Systems with Applications*, vol. 55, pp. 59–69, 2016.

[21] F. G. Mohammadi and H. Sajedi, "Region based image steganalysis using artificial bee colony," *Journal of Visual Communication and Image Representation*, vol. 44, pp. 214–226, 2017.

[22] G. Gando, T. Yamada, H. Sato, S. Oyama, and M. Kurihara, "Fine-tuning deep convolutional neural networks for distinguishing illustrations from photographs," *Expert Systems with Applications*, vol. 66, pp. 295–301, 2016.

[23] N. Rahmouni, V. Nozick, J. Yamagishi, and I. Echizen, "Distinguishing computer graphics from natural images using convolution neural networks," in *2017 IEEE workshop on information forensics and security (WIFS)*. IEEE, 2017, pp. 1–6.