



International Journal of Electronic Security and Digital Forensics

ISSN online: 1751-9128 - ISSN print: 1751-911X
<https://www.inderscience.com/ijesdf>

Application of artificial intelligence: methods to detect the image changes in social media

Yan Gao

DOI: [10.1504/IJESDF.2027.10070134](https://doi.org/10.1504/IJESDF.2027.10070134)

Article History:

Received:	02 September 2024
Last revised:	14 November 2024
Accepted:	20 November 2024
Published online:	13 April 2026

Application of artificial intelligence: methods to detect the image changes in social media

Yan Gao

Department of Media,
Zhengzhou University of Economics and Business,
Zhengzhou Henan, 451191, China
Email: gaoy09060430@163.com

Abstract: Active and passive methods are the two basic techniques to image modification detection that are described in the existing body of literature. In contrast to passive approaches, active techniques are proactive in nature, since they embed structures into photographs in order to enable future authenticity verification. The outputs from various networks are then fused together via the concatenation in order to determine whether or not the picture has been manipulated. This results in a complete detection framework that is more effective than the individual approaches that make up the framework. Our study presents a one-of-a-kind dataset that is the result of the combination of four datasets that are freely accessible to the public. This set of data includes photos that have been changed naturally and are very close to what happened in real life. It demonstrates improved generalisation across a wide range of data manipulation methods.

Keywords: digital image forensics; convolutional neural network; CNN; deep learning.

Reference to this paper should be made as follows: Gao, Y. (2026) 'Application of artificial intelligence: methods to detect the image changes in social media', *Int. J. Electronic Security and Digital Forensics*, Vol. 18, No. 8, pp.1–15.

Biographical notes: Yan Gao studied in Xinyang Normal University from 2002 to 2006 and received her Bachelor's degree in 2006. From 2007 to 2010, she studied in Guizhou Normal University and received her Master's degree in 2010. Currently, she works in Zhengzhou College of Economics and Business. She has published seven papers. Her research interests are included social media and national political development.

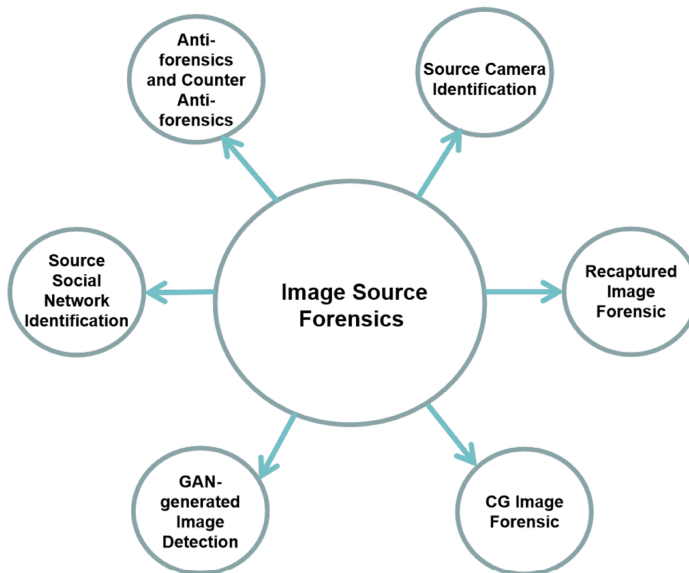
1 Introduction

Digital photography, which has become an important part of our daily lives, was made possible by the growth of modern technology. Software makes it easy to take digital pictures with our phones and digital reflex cameras. We can then share these pictures on social networking sites to let our friends know about important events in our lives. But it is also a bad thing that digital shots are so easy to work with. Every day, we see fake pictures more and more, and now we do not truly believe them (Zhu et al., 2004). This is even more important now that methods for artificial intelligence (AI) like generative

adversarial networks (GAN) (Farid, 2006) are being worked on. These tricks can be used by bad people to spread ‘fake news’ (Cao et al., 2019).

Over the past few years, many different methods (Beridze and Butcher, 2019), which are often grouped together as ‘digital image forensics’, have been developed to check if a digital picture is real. ‘Where is this digital image from?’ is a key question in this field of study (Piva, 2013). A field called source picture forensics tries to solve this question (Guan et al., 2019). There is a list of subtopics in Figure 1 that can be looked into to reach this goal. Imaging forensics with recorded images, imaging forensics with computer graphics (CG) (Hinton et al., 2006), GAN-generated image recognition (Barrios-Avilés et al., 2018), and imaging forensics with recorded images are some of these subtopics.

Figure 1 Review structure (see online version for colours)



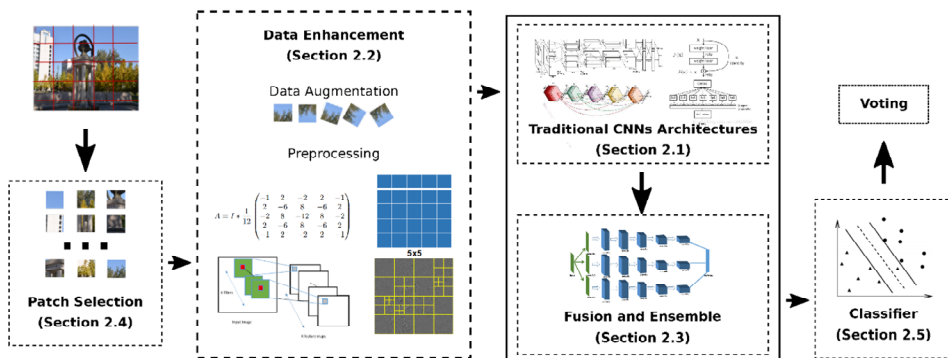
Source picture forensics used to be done with a number of methods that were based on statistics analysis and pattern recognition. Newer computers with more power have made people interested in methods based on machine learning all over again. To be more specific, systems based on deep learning have been used successfully in source picture forensics and have been shown to work well in a number of competitions (Bengio et al., 2007).

There are a lot of different names for techniques that use artificial neural networks. They have been around for a long time. Deep learning is a field that grew out of a set of important discoveries made in 2006 (Ranzato et al., 2007; LeCun et al., 2015; Goodfellow et al., 2016). Because of these findings, deep neural networks could be used, and a whole field was born (Khan et al., 2019) From then on, deep learning-based techniques have consistently done better than other methods based on more standard plans on a wide range of tasks, such as handwriting recognition and picture classification. There are a lot of different ways that have been made in the area of computer vision. Some of these ways are GAN, convolutional neural networks (CNN), and recurrent

neural networks (RNN). CNNs are one type that does well with jobs that involve images. These are the building blocks of many digital picture tracking systems because of this.

CNNs are made up of three main parts: activation functions, pooling layers, and convolutional layers. These layers are put together to make CNNs. To make CNN pattern, these pieces are stacked on top of each other. Khan et al. say that the most recent improvements in CNN built and trained can be broken down into four groups. You can put these things into groups like regularisation, loss function, structure reformulation, and parameter optimisations. All of them can make things work better, but framework reformulation is the most important. Space exploitation, depth, multi-path, breadth, channel boosting, feature map exploitation, and attention-based CNNs are the seven main groups it can be put into. A lot of people know that CNN systems like ResNet, DenseNet (Huang et al., 2017), Xception (Chopra et al., 2005), SENet (Chollet, 2017), and Siamese network (Jie et al., 2017) are popular. This review by Khan et al. (2019) and Guan et al. (2019) can help someone learn more about deep learning.

Figure 2 Framework for deep learning-based source forensics methods (see online version for colours)



Note: The related approach is detailed in Section 1.

Deep learning-based techniques in image source forensics have shown excellent results, which is what motivates us to present a full analysis of those methods. Our goal is to do so in a manner that will make it possible for a novice to enter this area with some assistance. For the purpose of this survey article, we made every effort to compile all of the relevant papers that have been published in arXiv, journals, and conferences. It has come to our attention that all of these schemes share some modules, and as a result, we suggest a unifying framework that can accommodate all of these schemes, therefore making it easier to compare them. A graphical representation of the framework may be seen in Figure 2. Both of these methods are viable options for achieving classification of a single patch. A dotted contour is used to indicate the optional pieces that are shown in Figure 2. Beginning with the overarching structure, we have separated the algorithms that were examined into the five distinct sub-sections that are listed below, according to the primary contributions that each algorithm made using regular CNN to find the source camera; performance improvement through data enhancement (DA), which includes both data addition and data preparation; a performance boost brought about by fusion and ensemble (F/E], the improvement of speed through patch selection (PS) adoption of many different categorisation schemes (C).

Tables 1 and 2 show an overview of the network factors that were thought to be the most important. After that, we made Tables 3 and 4 to show an overview of the test conditions and the results that different designs and some popular CNNs achieved.

The following table provides a summary of the architectures that were examined for source camera identification. Constrained convolutional operations are written as CC, whereas instant convolutional operations are written as IC. GAP is an abbreviation that indicates the existence of a global average mining operation. It is indicated by the symbol BN that batch normalisation procedures are present.

2 Literature review

The most advanced approaches for detecting alterations made to digital photos have been developed as a consequence of recent developments in image forensic techniques. These approaches are regarded as cutting-edge techniques. Previous research investigations have provided techniques based on observations made at each stage of the image history, starting with the phase of capture and concluding with the phase of storing the image in a compressed format. A trace is left on the picture as a result of the processing of the image, which may be used to verify the digital authenticity of the image (Krizhevsky et al., 2012). After that, the verification of a digital signature is used to verify whether or not it is legitimate.

This method does not include the installation of digital watermarks in the photographs, nor does it involve the comparison of the images for the purposes of training and testing. It was claimed by the authors that the individual picture characteristics that were retrieved during the acquisition phase represent evidence in and of themselves that the image is genuine (Simonyan and Zisserman, 2014). In many cases, these characteristics may be seen with the naked eye. In particular, it employs visual artefacts that are brought about by a variety of anomalies as markers in order to ascertain the authenticity of the image developed a method that makes use of a colour filter array in order to identify instances of picture manipulation (Schaefer and Stich, 2013). An individual feature and a straightforward threshold-based classifier are both computed by it. Image manipulation, computer-generated pictures, and genuine photos were used by the authors to test their methodology. In the experimental investigation, the error rates were shown to be minimal.

The first branch looks for areas that can be changed by using visual effects. The second branch, on the other hand, uses visual patterns to look for areas that can become copies or moved. For BusterNet training to work well, they suggested using simple methods and a step-by-step process for datasets that were not related to the study. Based on their in-depth study, BusterNet did much better than regular copy/move algorithms by a large and significant range. Our study used the CASIA and CoMoFoD datasets to test the suggested design.

Table 1 Summary of the reviewed architectures for source camera identification

Architecture	Input size	Preprocessing	Convolutional part			Fully connected part				
			N layers	Pooling	Activation	BN	GAP	N layers	Activation	Dropout
A1	$48 \times 48 \times 3$	-	3	ReLU	Max	-	-	1	ReLU	✓
A2	$32 \times 32 \times 3$	-	2	L-ReLU	Max	-	-	2	L-ReLU	✓
A3	$36 \times 36 \times 3$	-	3	ReLU	Avg	✓	-	1	ReLU	✓
A4	$64 \times 64 \times 3$	-	13	ReLU	Max	-	-	2	?	✓
A4	$256 \times 256 \times 3$	-	1 conv, 12 residual	ReLU	-	-	✓	-	-	1 conv, 12 residual

Notes: CC and IC denote constrained and instant convolutional operations, respectively. GAP denotes the presence of a global average pooling operation. BN denotes the presence of batch normalisation operations.

They discussed the approach of passive picture forensic analysis and brought attention to the obstacles that lie ahead in the process of building a system for the identification of images that have been altered. In a different piece of research (Amerini et al., 2011), I developed a novel IDF method that was supported by a CNN. One of the objectives of this method is to acquire an automated understanding of how picture alteration might be carried out. As input, the image-altering characteristics that are formed after the contents of a picture are destroyed are used by the IDF approach that has been discussed. These methods focus on looking at the local operational link between pixels instead of the look and feel of the picture because messing with them could change some resident relationships. It can then find examples of fakes in a picture. Finding fake digital pictures was the subject of another study, which recommended using a CNN-based structure for the process. Their argument was that the preparation step is closely linked to the CNN model's first layer. It searches for problems that happen because of manipulation in this step (Bayar and Stamm, 2016). They utilised trial photos to train the CNN model, while the support vector machine (SVM) was employed to identify any alterations that were present. A ringed residual U-Net, sometimes known as an RRU-Net, was suggested for the purpose of detecting forgeries in picture slicing. They suggested an architecture that makes use of an end-to-end image segmentation network in order to identify instances of counterfeiting with increased accuracy.

In a different study, Rao and Ni (2016) offered a method based on transfer learning that utilises the steganalysis model to uncover previously unknown data. When applied to the BOSSBase and BOW datasets, this method gave them an average accuracy rate of 97.36%. An approach based on transfer learning was shown in. Using the AlexNet model's pre-trained weights in this process cuts down on the time needed for training. A predictor in this method is the SVM. Generally, the model did a good job.

The network that the authors showed can do many jobs at once and has multiple output streams. In this case, the surface label will be obtained by one of these streams, and the edge of the interface region will be obtained by the stream that follows (Bi et al., 2019). A innovative approach to photo splicing detection using a feature-based algorithm was reported in the publication (Zhan et al., 2017). This approach use the combination of photos that occur together in order to compute local characteristics (Doegar and Gaurav, 2019). After that, the feature parameters are retrieved using the local features (Salloum et al., 2018). By combining the segmentation process with the expectation-maximisation approach, learning is achieved. This is because photos that have been spliced and hosted may exhibit unique features.

Based on the study that was shown earlier, most of the methods used to find fakes are based on manual methods for extracting features. The person doing the work is very important for these methods to work (Cozzolino et al., 2015). Deep learning-based methods have led to the creation of an automated feature extraction system. Therefore, using deep learning takes away the chance of human error, Bishop (1995) making the model more accurate and reducing the time needed to complete the process.

3 Materials and methods

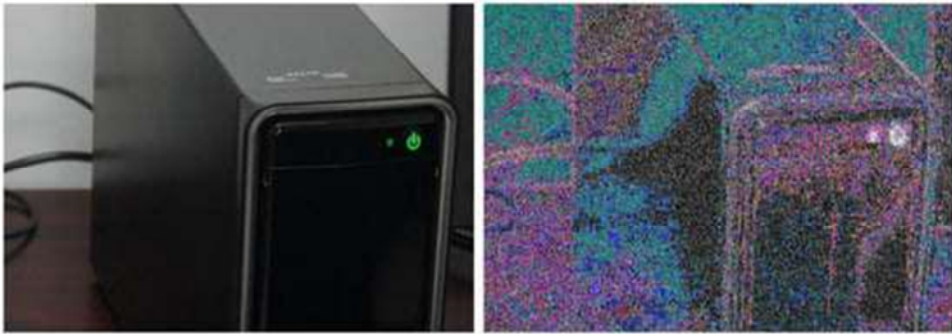
This section will talk about the method that was used to find images that had been changed. In addition to a new and innovative multi-stream neural network design, this method uses error-level analysis (ELA) and discrete wavelet transform (DWT).

It was because we learnt that traditional passive methods can help a CNN learn how to identify picture changes that we decided to use a multi-stream neural network design. We did not try to figure out which way is the best overall; instead, we looked at this option as a first step towards finding out what benefits such a system might offer. From this point of view, we wanted to get different traits out of the picture data, so we chose two streams based on methods described in other research. We came up with this way because we wanted to use some passive image analysis techniques in our research into how well multi-stream systems can find changed pictures. It was an important turning point in our study.

3.1 Error-level analysis

The process is shown in Figure 3, and the code used to make the pictures can be found in Listing 1. In the past, human forensics experts have used ELA to make changes in JPEG pictures' format flaws stand out more. It works by comparing a JPEG picture with different levels of quality. Any change in compression rate is then made to stand out more on the new image.

Figure 3 ELA example (see online version for colours)



Note: Transformed power button image (left). ELA image, mostly black except for two compressed parts.

3.2 DWLT

'DWT' stands for 'denoising', which is a scientific way to look at signs that can be used to make shots better. It takes a picture and breaks it up into a group of frequency components. Each frequency component shows a different amount of sharpness or detail. People often use DWT to pick out parts of a picture in order to find changes. In this case, it works well to reduce the size of picture data while keeping important image qualities. A picture is broken up into four smaller pieces, or 'sub-bands', each with a different frequency. The DWT algorithm can then do its job. This part is for material with a low frequency. For high-frequency data, there are parts called low-high (LH), high-low (HL), and high-high (HH). The equations (1) and (2) are used to do isomorphic DWT. How to do inverse DWT is shown in equation (3).

$$W_\varphi(j_0, k) = \frac{1}{\sqrt{M}} \sum_M f(x) \varphi_{j_0, k}(x) \quad (1)$$

$$f(x) = \frac{1}{\sqrt{M}} \sum_k W_\varphi(j_0, k) \varphi_{j_0, k}(x) \quad (2)$$

The methodology suggested therein is used in this study and will be called the DWT method from here on. This method eliminates less significant characteristics from a picture by using the Laplace operator, bilateral filters, and DWT. As a consequence, the final picture alone has features with noticeable pixel variation – a frequent sign of fraud.

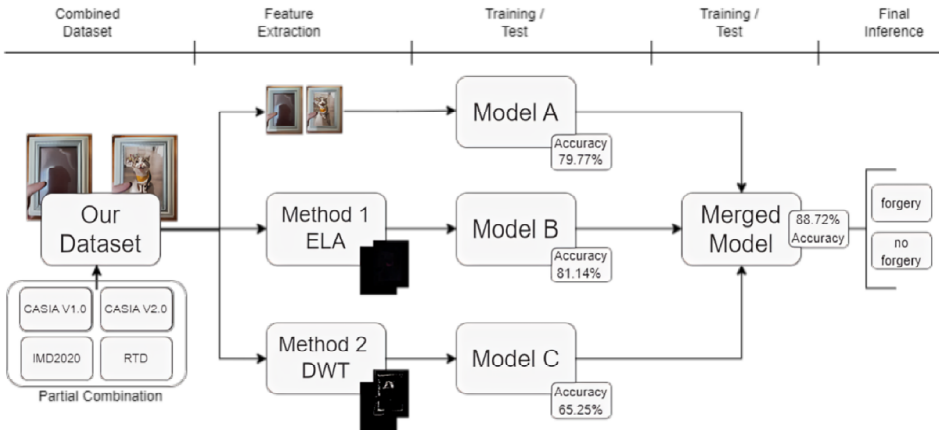
First, the picture is turned into greyscale. Next, DWT is used to separate the image data into sub-bands. After removing the LL band and a bidirectional filter, the image is put back together again. Finally, the Laplace operator is used. It is easier to see the sharp changes between pixels this way, but this method does not show the changed areas if filters are used to hide them.

3.3 Proposed method

Beginning with the ELA method and the method recommended by this method is what is being shown. This is how two more sets of pictures are made as part of the feature selection step. These sets will be used as different lines for the next model architecture.

More learning layers are added, and the model as a whole is changed while it is being trained to get to the level of accuracy that is wanted. See Figure 4 for a visual picture of these stages. This image also shows the overall structure of the system.

Figure 4 Illustration of the proposed approach (see online version for colours)

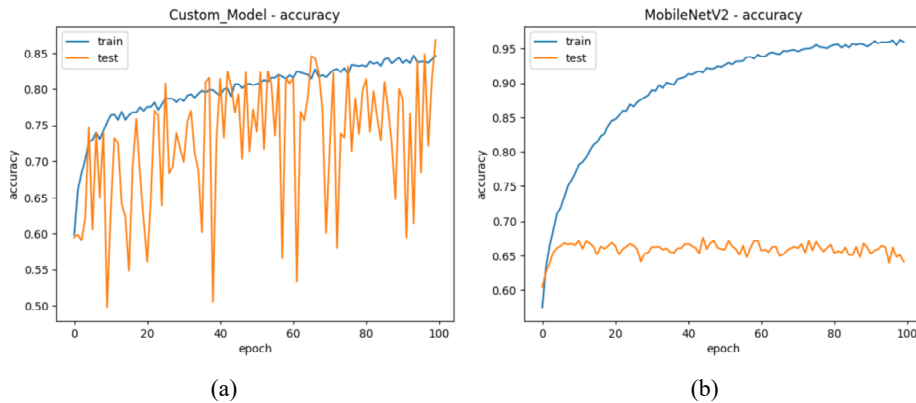


The purpose of this strategy is to investigate the accuracy of each stream separately in addition to the accuracy of the combination of streams. In order to do this, we trained three separate models on our own, as well as a model that included all three streams, in addition to several extra learning layers. Photo entry for stream A is limited to original shots that have not been changed in any way. Image B only uses pictures picked out for their features using ELA, and image C only uses pictures picked out using the DWT

method. There are four levels added to the merged model after the models are put together. These three sets of pictures are put into this model.

We looked into using models that had already been trained when we first started to think about using passive methods in our work. We chose the MobileNetV2 pre-trained model because it is a lightweight one. Figure 5 (which you can find here) shows the results of this analysis.

Figure 5 (a) Custom model accuracy vs. MobileNetV2 (b) Custom model training accuracy training MobileNetV2 accuracy beforehand (see online version for colours)



3.4 Dataset assembly

Our studies began with the compilation of the final dataset, which was the first stage in the execution process. In order to do this, we painstakingly chose genuine and modified photographs from four different databases. Due to the fact that these datasets required human alteration of photographs, with a focus on making changes that were difficult to detect, they were selected because they provided a reliable portrayal of real-life circumstances. The size of the dataset was restricted since there was a limited supply of photographs that had been modified by humans. The creation of these images requires a substantial amount of time and effort.

CASIA V2.0, which was recommended in, has 7491 real pictures and 5,123 changed images that have been spliced, duplicated, and retouched on top of mask changes.

CASIA Version 1.0: Proposed by There are a total of 462 splicing photos, 800 legitimate photographs, and 459 duplicate-type modification images included in this collection. None of the retouching procedures have been done to this dataset.

We separated the photographs into two files so that the dataset would be more balanced and so that we could include images that had been realistically modified. There are 7,491 genuine photos taken from the CASIA V2.0 dataset that are included inside the first folder. There are 7,491 changed photos in the second folder. It was done this way to make sure that the number of edited and unaltered pictures was the same. Yet, it is not known if 50% of the time, applications used in the real world are manipulated, which could lead to classifier bias. Table 2 shows an example of how the pictures from each dataset are spread out.

Table 2 Our dataset uses photos from each dataset

	<i>CASIA V1.0</i>	<i>CASIA V2.0</i>	<i>IMD2020</i>	<i>RTD</i>	<i>Total</i>
Authentic	0	6,859	0	0	6,859
Tampered	274	4,289	2,023	220	7,491

The end collection that was used in this project has a total of 14,982 photos, with half being the originals and the other half having been changed. All of the photos were changed to .jpg file before they were handled because some of the images could not be used by the model.

4 Experimental results

PyTorch was used to carry out our methods, and the datasets we talked about in Section 3.1 were used to test them. In the next part, the testing results from both the larger dataset and the smaller dataset are shown, along with some similarities.

4.1 Results on full datasets

The very first thing we did was run tests on all of the information. You can choose from four full files, and each one has pictures made by a computer using one of four powerful graphics engines: Artlantis, Autodesk, Corona, or V-Ray. In Section 3.1, you can find this information. There were a few different ways we tested the neural network’s investigative performance while it was being trained on each of the four datasets. These included adding more data and/or changing the loss function. We tested and reported the classification success of each network on four test sets. Each network was trained on a different dataset, like Autodesk’s training set. In these test sets, there were CG pictures from four different rendering engines that were not used during the training phase.

Table 3 shows the findings. The four full datasets from Artlantis, Autodesk, Corona, and V-Ray were used to train the network and get those results. It gives individual scores and the average of all the test sets. It also shows how well the different types of questions worked. The tables with italics show how well the assumption works on test sets that have pictures made by computers from rendering engines that are not known. We show the results of normal training, which is the same as the baseline method, training with noise addition (NA) augmentation, training with Gaussian blurring (GB) augmentation, training with colour jitter (CJ) augmentation, training with colour transfer (CT) augmentation, training with the modified new loss that takes both global and local predictions into account, training with the augmentation of GB + CJ, training with the augmentation of GB + CT, and training with the modified new loss and the augmentation of GB + CT.

The results in Tables 3 and 4 are interesting, and we need to talk about them. The last column of each of the four tables shows that GB did a lot better than noise addition (NA) in the first group of improvement methods. This is because the average classification accuracy was much higher. Overall, GB was more accurate on all four datasets than it was during normal training. However, NA could only make things better when it was trained on Corona. Noise addition, which adds new high-frequency parts to try to ‘cover’ the old ones, is not as good as cleaning, which takes out some of the high-frequency

parts. Because of this, smoothness might be a better way to lessen the effect of how the training images were handled in the past.

Table 3 The average test accuracy (last column, the major performance parameter) using networks trained on Artlantis' whole training set and experimental test accuracy on the four test sets

<i>Tested on methods</i>	<i>Artlantis</i>	<i>Autodesk</i>	<i>Corona</i>	<i>VRay</i>	<i>Average</i>
The usual training	98.69%	89.94%	85.42%	88.14%	87.65%
With aug. NA	98.75%	81.25%	79.31%	89.44%	87.19%
With aug. GB	98.75%	91.39%	90.00%	94.17%	93.58%
With aug. CJ	98.89%	88.06%	87.92%	92.36%	91.81%

Notes: There are italicised generalisations. Here, 'aug.' denotes augmentation. Denotes highest average test accuracy.

Table 4 Average test accuracy (final column, the major performance parameter) for networks trained on the whole Autodesk training set for the four test sets

<i>Tested on methods</i>	<i>Artlantis</i>	<i>Autodesk</i>	<i>Corona</i>	<i>VRay</i>	<i>Average</i>
Normal training	90.61%	98.44%	85.42%	88.14%	92.36%
Along with August NA	95.56%	98.33%	79.31%	89.44%	87.19%
Utilising August GB	89.17%	98.33%	90.00%	94.17%	93.58%
With August CJ	90.69%	98.98%	87.92%	92.36%	91.81%

Notes: There are italicised generalisations. Here, 'aug.' denotes augmentation. Denotes highest average test accuracy.

To begin, we discover that GB and CT work better together than GB and CJ when we combine them. No matter what, this is always the case. CJ does better than CT on its own when taught on Artlantis, Autodesk, and Corona, but not when taught on VRay. This is a very interesting event. GB is trying to get rid of some of the processing history records that CJ adds random noise to pictures. This may make them a little bigger. This means that using GB and CJ together might make one of them less useful than using it alone. This is one reason why what happened might have happened. The GB+CT mix did the best on average tests of all the methods that were trained on Artlantis. As it stands, both GB and CT can meet their goals of making training samples more diverse and less dependent on the processing past. They will not be able to stop each other. Finally, getting good results in all four tables is also possible when you use modified loss, GB, and CT together. In fact, this combination gets the best general test score (97.19%).

4.2 Comparisons of test accuracy and training time

These methods also take into consideration the difficult circumstances of generalisation performances and the shortage of training data. Comparisons are made with regard to the correctness of the exam and the amount of computing resources required for training.

First, we compare our method to the one that said an extra, longer training step could improve the success of generalisation. looks at how our better training method and the different forms of our approach change when they are trained on the whole Artlantis training set and how much it takes to run each method. One NVIDIA RTX A6000 graphics processing unit (GPU) was used for all the tests. When we compare our different

training methods to regular training, as shown in Table 5, the total time spent training goes up by about 10 to 20 minutes. It does not take more than ten minutes longer to train with the suggested changes to the loss function and network design, as well as the ways to make things better. It takes about 23 milliseconds for all learnt networks that were trained in different ways to draw a conclusion from each test picture. These include standard training, extra-improved training, our different data-augmentation methods, our new loss, and any mix of the methods we've mentioned. This is the way all learnt networks work. In this way, the test speed is kept high, and the 43 test shots are taken as quickly as possible. One pair of test accuracy numbers is shown between each other in Table 5 as a '/'. The first pair shows the result, and the second pair shows our method. Our method works just as well as, or even a little better than, what is thought to be the most up-to-date way to do research tests. Another thing is that our method always gets better test results overall, across all four levels of testing. Along these lines, it is important to note that our method works just as well as the better training method while using a lot less computer power for training (Table 5).

Table 5 Different approaches' network training times (in minutes) for the whole Artlantis training set

<i>Methods</i>	<i>Training time</i>	<i>Additional time compared to normal training</i>
The usual training	480	-
With more and better training	782	+435
Along with aug. GB	356	+9
With augment. GB + CT	356	+9
With a new loss	355	+8
With a new loss and an increase of GB + CT	368	+21

4.3 Discussion

The earlier tests show that our suggested ways of improving data and the changed new loss function can make the CG forensics problem work better when used together in different test settings. In research studies, we can get better results when we mix these methods from different groups. You can get the best total performance when you use the adjusted new loss along with the GB and CT processes. In particular, this lets you get the best average test scores when things are hard and there are not many training examples. In this case, doing these two things together gives us the most accurate results. Our suggested methods, which are meant to achieve various goals (e.g., lowering the effect of past data processing, raising the variety of data, and learning all-encompassing features by combining local and global predictions), seem to be able to work well together. This is because each of our methods is meant to reach a different goal. Our ways are also easier to understand theoretically and faster to run on computers, as shown by the comparisons with other methods that are already in use.

5 Conclusions

We suggested using efficient yet effective methods that include the right boosting operations, a slightly changed network design, and a loss function. The goal was to learn more useful features that would help us deal with the tough situations we were thinking about. In contrast, the methods that were already used are still very computationally expensive during the training phase. Several tests were done on sets of computer-generated pictures (CG) made by four different advanced graphics rendering systems. We did tests that showed our suggested methods can help improve the performance of forensic analysis. Specifically, they can help with generalisation and classification when there are not many training samples to use. For picture analysis methods to work better in the real world, where they are likely to be put to use in the tough situations we talked about earlier, they should have these good qualities. In terms of classification accuracy and computing speed during the training phase, our methods do better than the ‘state-of-the-art’ methods, as shown by the comparisons of experiments.

In the years to come, our work could get better in a lot of ways. Additionally, we will carefully examine the spectral and statistical features of NIs and CG images to find out more about the variations between the two types of images and to determine what kind of improvement operation would help the detective work go more smoothly. We should make our own operation instead of using the usual Gaussian filtering operation before we start studying the boosting operation. As a result, the way data was handled in the past would have less of an impact. Secondly, we want to know if it is possible to move higher-order colour factors from one picture to another so that we have more training data. This can be done by making the normal CT method bigger. This method considers both first- and second-order traits. Lastly, it would be interesting to see how the number of local guesses changes things and if there is a way to mix world and local data in a more advanced way. Knowing the similarities and differences between drawing tools would be helpful in the future if you can. The goal is to figure out and explain the performance drops, which usually come in different amounts when it comes to being able to use forensically across CG images that were made with various tools. Additionally, we wish to improve our methods so they can be applied to more video security issues.

References

- Amerini, I., Ballan, L., Caldelli, R., Del Bimbo, A. and Serra, G. (2011) ‘A SIFT-based forensic method for copy-move attack detection and transformation recovery’, *IEEE Trans. Inf. Forensics Secur.*, Vol. 6, pp.1099–1110.
- Barrios-Avilés, J. et al. (2018) ‘Movement detection with event-based cameras: comparison with frame-based cameras in robot object tracking using powerlink communication’, *Electronics*, Vol. 7, No. 11, p.304.
- Bayar, B. and Stamm, M.C. (2016) ‘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*, Vigo, Spain, 20–22 June, pp.5–10.
- Bengio, Y., Lamblin, P., Popovici, D. and Larochelle, H. (2007) ‘Greedy layer-wise training of deep networks’, in *Advances in Neural Information Processing Systems*, pp.153–160, MIT Press, Cambridge, MA, USA.
- Beridze, I. and Butcher, J. (2019) ‘When seeing is no longer believing’, *Nat. Mach. Intell.*, Vol. 1, pp.332–334.

- Bi, X., Wei, Y., Xiao, B. and Li, W. (2019) 'RRU-Net: the ringed residual U-Net for image splicing forgery detection', in *Proceedings of the 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, Long Beach, CA, USA, 16–17 June, pp.30–39.
- Bishop, C.M. (1995) *Neural Networks for Pattern Recognition*, Oxford University Press, Oxford, UK, 1995.
- Cao, Y.J., Jia, L.L., Chen, Y.X., Lin, N., Yang, C., Zhang, B., Liu, Z., Li, X.X. and Dai, H.H. (2019) 'Recent advances of generative adversarial networks in computer vision', *IEEE Access*, Vol. 7, pp.14985–15006.
- Chollet, F. (2017) 'Xception: deep learning with depthwise separable convolutions', in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Honolulu, HI, USA, 21–26 July, pp.1251–1258.
- Chopra, S., Hadsell, R. and LeCun, Y. (2005) 'Learning a similarity metric discriminatively, with application to face verification', in *Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*, Vol. 1, IEEE Computer Society, Washington, DC, USA, pp.539–546.
- Cozzolino, D., Poggi, G. and Verdoliva, L. (2015) 'Splicebuster: a new blind image splicing detector', in *Proceedings of the 2015 IEEE International Workshop on Information Forensics and Security (WIFS)*, Rome, Italy, 16–19 November, pp.1–6.
- Doegar, M.D.A. and Gaurav, K. (2019) 'CNN based image forgery detection using pre-trained AlexNet model', *Int. J. Comput. Intell. IoT 2019*, Vol. 2, p.6.
- Farid, H. (2006) 'Digital doctoring: how to tell the real from the fake', *Significance*, Vol. 3, pp.162–166.
- Goodfellow, I., Bengio, Y. and Courville, A. (2016) *Deep Learning*, MIT Press, Cambridge, MA, USA.
- Guan, H., Kozak, M., Robertson, E., Lee, Y., Yates, A.N., Delgado, A., Zhou, D., Kheyrkhan, T., Smith, J. and Fiscus, J. (2019) 'MFC datasets: large-scale benchmark datasets for media forensic challenge evaluation', in *Proceedings of the 2019 IEEE Winter Applications of Computer Vision Workshops (WACVW)*, Waikoloa Village, HI, USA, 7–11 January, pp.63–72.
- Hinton, G.E., Osindero, S. and Teh, Y.W. (2006) 'A fast learning algorithm for deep belief nets', *Neural Comput.*, Vol. 18, No. 7, pp.1527–1554.
- Huang, G., Liu, Z., Van Der Maaten, L. and Weinberger, K.Q. (2017) 'Densely connected convolutional networks', in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Honolulu, HI, USA, 21–26 July, pp.4700–4708.
- Jie, H., Li, S., Albanie, S., Gang, S. and Wu, E. (2017) *Squeeze-and-Excitation Networks*, arXiv 2017, arXiv:1709.01507.
- Khan, A., Sohail, A., Zahoora, U. and Qureshi, A.S. (2019) *A Survey of the Recent Architectures of Deep Convolutional Neural Networks*, arXiv, arXiv:1901.06032.
- Krizhevsky, A., Sutskever, I. and Hinton, G.E. (2012) *Advances in Neural Information Processing Systems*, pp.1097–1105, MIT Press, Cambridge, MA, USA.
- LeCun, Y., Bengio, Y. and Hinton, G. (2015) 'Deep learning', *Nature*, Vol. 521, p.436.
- Piva, A. (2013) 'An overview on image forensics', *ISRN Signal Process.*
- Ranzato, M., Poultney, C., Chopra, S. and LeCun, Y. (2007) 'Efficient learning of sparse representations with an energy-based model', in *Advances in Neural Information Processing Systems*, pp.1137–1144, MIT Press, Cambridge, MA, USA.
- Rao, Y. and Ni, J. (2016) '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)*, Abu Dhabi, United Arab Emirates, 4–7 December, pp.1–6.
- Salloum, R., Ren, Y. and Kuo, C-C.J. (2018) 'Image splicing localization using a multi-task fully convolutional network (MFCN)', *J. Vis. Commun. Image Represent.*, Vol. 51, pp.201–209.

- Schaefer, G. and Stich, M. (2003) 'UCID: an uncompressed color image database', *Storage Retr. Methods Appl. Multimed.*, Vol. 5307, No. 12, pp.472–480.
- Simonyan, K. and Zisserman, A. (2014) *Very Deep Convolutional Networks for Large-Scale Image Recognition*, arXiv, arXiv:1409.1556.
- Zhan, Y., Chen, Y., Zhang, Q. and Kang, X. (2017) 'Image forensics based on transfer learning and convolutional neural network', in *Proceedings of the 5th ACM Workshop on Information Hiding and Multimedia Security*, Philadelphia, PA, USA, 20–22 June.
- Zhu, B.B., Swanson, M.D. and Tewfik, A.H. (2004) 'When seeing isn't believing [multimedia authentication technologies]', *IEEE Signal Process. Mag.*, Vol. 21, pp.40–49.