applied sciences

MDPI

Article

Deep Learning-Based Digital Image Forgery Detection System

Emad Ul Haq Qazi *, Tanveer Zia and Abdulrazaq Almorjan

Center of Excellence in Cybercrimes and Digital Forensics (CoECDF), Naif Arab University for Security Sciences (NAUSS), Riyadh 14812, Saudi Arabia; tzia@nauss.edu.sa (T.Z.); t-aalmargan@nauss.edu.sa (A.A.)
* Correspondence: qabdulrab@nauss.edu.sa

Abstract: The advancements of technology in every aspect of the current age are leading to the misuse of data. Researchers, therefore, face the challenging task of identifying these manipulated forms of data and distinguishing the real data from the manipulated. Splicing is one of the most common techniques used for digital image tampering; a selected area copied from the same or another image is pasted in an image. Image forgery detection is considered a reliable way to verify the authenticity of digital images. In this study, we proposed an approach based on the state-of-the-art deep learning architecture of ResNet50v2. The proposed model takes image batches as input and utilizes the weights of a YOLO convolutional neural network (CNN) by using the architecture of ResNet50v2. In this study, we used the CASIA_v1 and CASIA_v2 benchmark datasets, which contain two distinct categories, original and forgery, to detect image splicing. We used 80% of the data for the training and the remaining 20% for testing purposes. We also performed a comparative analysis between existing approaches and our proposed system. We evaluated the performance of our technique with the CASIA_v1 and CASIA_v2 datasets. Since the CASIA_v2 dataset is more comprehensive compared to the CASIA_v1 dataset, we obtained 99.3% accuracy for the fine-tuned model using transfer learning and 81% accuracy without transfer learning with the CASIA_v2 dataset. The results show the superiority of the proposed system.

Keywords: machine learning; deep learning; image forgery; ResNet50; YOLO CNN; CASIA

check for updates

Citation: Qazi, E.U.H.; Zia, T.; Almorjan, A. Deep Learning-Based Digital Image Forgery Detection System. *Appl. Sci.* **2022**, *12*, 2851. https://doi.org/10.3390/app12062851

Academic Editor: David Megías

Received: 11 November 2021
Accepted: 9 December 2021
Published: 10 March 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Digital images have an important role in many fields such as in newspapers, digital forensics, scientific research, medicine, and so forth. Nowadays, the usage and sharing of digital images on social media platforms is also widespread. Digital images are considered one of the main sources of information. Considering the excessive use of image sharing through various social media platforms such as WhatsApp, Instagram, Telegram, and Reddit, differentiating between real and forged images is a challenging task. The availability of many image editing software applications is making it more difficult to detect the authenticity of an image day by day. There are generally two approaches that image manipulation can be categorized into, as follows:

1. Active approach;
2. Passive approach.

With the active approach, a watermark or digital signature is embedded when the image is created. While using these embeddings, whether the image has been tampered with or not is analyzed at later stages.

In the passive approach, any pre-embedded information, such as a watermark embedded for the detection of image forgery, cannot be relied upon. This approach is also known as the blind approach because there is no additional information for image forgery detection. This approach is based on features that are extracted directly from the images.

Furthermore, the passive approach can be categorized into two types—independent and dependent. The independent approach detects resampling and compression forgeries, while the dependent approach detects splicing and copy/move forgeries. Figure 1 shows the hierarchy of the above approaches.

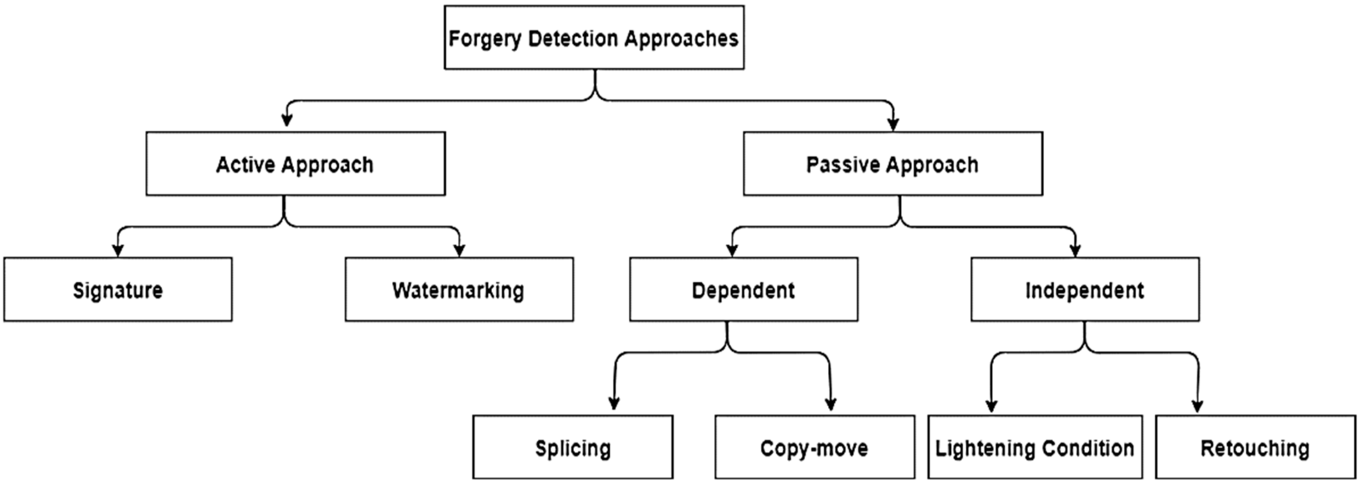

Figure 1. Forgery detection hierarchy.

In the copy/move type of manipulation, a particular part of an image is selected and then copied and pasted to another part of the same image. By doing this, the correlation value of these two parts of the image will be relatively higher compared to other parts of the image. The goal of copy/move forgery detection is to correctly identify these duplicates in these images by comparison of the attributes extracted from the features using distance measures. Two approaches, as follows, are commonly used to extract patch-wise features from the images:

1. Images are divided into blocks, and features are extracted from these blocks, as discussed in [1];
2. Key points are identified from the image and these key point features are extracted.

The features extracted in blocks or key points are compared one by one for generating matched pairs. The duplication is confirmed if matching is found among two blocks; it confirms the duplication and can be categorized as a manipulated image. The steps of the process are shown in Figure 2. Digital image splicing is a method of extracting of objects from one image and inserting those objects into another image.

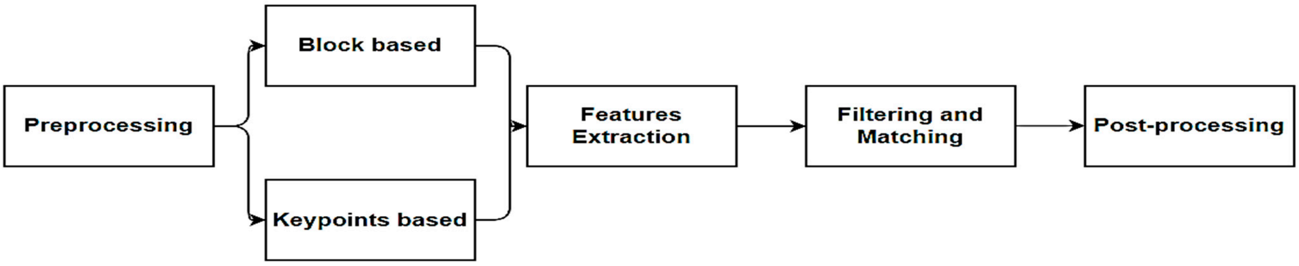

Figure 2. Copy/move detection.

It is easier to detect manipulations in copy/move image forgery detection compared to image splicing because the similar contours of an object can be easily detected in the same image since they have the same sizes, transitions, and textures. Different objects are introduced in the case of image splicing, with different textures, sizes, and transition attributes, and this approach makes forgery difficult to identify [2].

Image splicing forgery detection is dependent on the clues that are left after the manipulation of images. Some common image splicing issues include inconsistency, edge discontinuity because of the camera, and geometric and lighting conditions. Capturing an image from different cameras results in different attributes, and clue tampering can be confirmed [3]. There can be lighting inconsistencies as well, which can arise due to different lighting conditions. A double quantization effect can also arise when saving JPEG images because of two consecutive compression operations that are performed on the tampered image [2,3].

Image tampering usually does not have any visible clues though which one can tell whether the image has been tampered with or not; however, some statistics of the image may be altered. Christlein et al. [4] experimented on the copy/move approach and cut/paste-based detection methodology was discussed by Zambpglou et al. [5]; some of these approaches are shown in Figure 1.

The development of deep learning has led to improving methodologies where state-of-the-art methods, such as CNN, Mobile Net, and ResNet50v2, automatically extract the potential features, having been trained on large datasets. Some of the examples of CNN-based feature extractions are deep features utilized for image quality assessment [6], skin lesion classification [7], or person re-identification [8]. These extracted features are adapted into the inherent structural patterns of the data. This is the main reason behind their non-discriminative and robust architecture compared to the hand-engineered features.

In this paper, motivated by the deep learning technique, we propose a transfer learning-based approach. It is an effective architecture with which we incorporated the weights of a model previously trained on a large database, and hence, it benefitted from using the meaningful weights without having to train the model from scratch. We present an architecture based on the ResNet50v2 architecture that employs the use of transfer learning for the detection of tampered images, specifically, spliced images. We used the pre-trained weights of a YOLO CNN model to detect images that were specifically tampered with using the image splicing technique. Furthermore, this study makes the following contributions to this field of research:

- Detailed analysis between deep learning and hand-engineered techniques;
- Proposed ResNet50v2-based architecture for the authentication of original and forged images;
- Utilization of the transfer learning technique to effectively train our proposed model on benchmark datasets CASIA_v1 and CASIA_v2 [9];
- Discussion of the limitations and future directions of this research that can be carried forward.

The rest of the paper is organized as follows. Section 2 discusses the literature review and Section 3 explains the proposed system architecture. In Section 4, we present the dataset details. Section 5 presents the experimental results, discussion, and future work. Section 6 presents the conclusion.

2. Literature Review

Recent developments of image forensic techniques have led to the emergence of state-of-the-art techniques with which we can detect manipulations that have been made in digital images. Previously, some research studies [10–12] have proposed approaches that rely on the observations that are made during each phase of the image history, from its acquisition phase to saving it in a compressed format. The processing of the image leaves a trace on the image for the verification of digital authenticity. It is then determined as authentic or inauthentic by the verification of a digital signature.

Yerushalmy et al. [2] suggested a new approach for the detection of image forgery. This technique is not adding digital watermarking in the images and does not compare the images for training and testing. The authors proposed that image features extracted during the acquisition phase are themselves proof of authenticity of the image. These features are often visible to the naked eye. Specifically, it uses image artifacts caused by various irregularities as markers to determine image validity. Ahmet et al. [3] proposed a technique for detecting image tampering using a color filter array. It computes a single feature and a simple threshold-based classifier. The authors tested their approach with authentic, computer-generated, and tampered images. The experimental analysis showed low error rates.

Barad et al. [13] performed a research survey that was based on deep learning techniques for the task of image forgery detection, and they presented an analysis of the approaches used to detect the authenticity of images on publicly available datasets. Yue et al. [14] introduced a deep learning-based architecture for copy/move image forgery detection using BusterNet, which is an end-to-end trainable approach. BusterNet uses two-branch architecture. The goal of the first branch is to identify manipulation areas using visual artifacts, whereas the second branch identifies copy/move areas using visual similarities. For effective BusterNet training, they proposed simple techniques for out-of-domain datasets and a stepwise approach. Their extensive research study demonstrated that BusterNet outperformed traditional copy/move algorithms by a large margin. The proposed architecture was evaluated with the CASIA and CoMoFoD datasets.

Manjunatha et al. [15] discussed the importance of detecting tampering in images using deep learning-based techniques on publicly available datasets such as CASIA, UCID, MICC [9,16,17], and so forth. They covered passive image forensic analysis methodology and highlighted future challenges in developing a mechanism for the detection of tampered images. In another study, Belhassen et al. [18] proposed a unique IDF technique based on a CNN. The goal of this technique is to automatically learn how image modification could be done. The proposed IDF technique takes image-altering features as input generated after destroying the contents of an image. Since tampering alters some resident associations, this technique focused on examining the local operational association among pixels rather than focusing on the look and feel of the image; it then detects forgery in an image. In another study, Rao et al. [19] proposed a CNN-based architecture for the detection of digital image forgery. They proposed that the first layer of the CNN model is directly involved in the preprocessing stage. It searches for the issues that occur after tampering. They trained the CNN model on trial images, whereas SVM was used for the detection of manipulations. Bi et al. [20] proposed a ringed residual U-Net (RRU-Net) for forgery detection in image slicing. They proposed an architecture where forgery detection is employed using an end-to-end image segmentation network. The goal of the RRU-Net study was to use human brain mechanisms to develop an approach using RRU-Nets, which can detect manipulations without pre- and post-processing. Generally, the human brain works on recall and consolidation mechanisms. Therefore, the purpose of this technique is to optimize the learning capacity of a CNN, which is inspired by human brain attributes. They solved the gradient degradation problem, as residual propagation is used to recall the input feature information in a CNN. Finally, it differentiates between the original and fake regions, as the remaining response is merged with the response feature. The experimental results showed that the proposed technique gave better results compared to the state-of-the-art traditional methods.

In another study, Zhan et al. [21] proposed a transfer learning-based methodology that has the benefit of gaining prior knowledge using the steganalysis model. With this approach, they were able to obtain an average accuracy of 97.36% on the BOSSBase and BOW datasets. Amit et al. [22] proposed a transfer learning-based mechanism that utilizes the pre-trained weights of the AlexNet model, which saves training time. This approach uses SVM as a classifier. The overall performance of the model was satisfactory.

Salloum et al. [23] suggested the use of a multitasking fully connected network. Since a single-task fully connected network has irregular output, the proposed technique performed better compared to the single-task fully connected network. The authors proposed a multitask fully connected network comprising a collection of output streams. One of these streams acquires the surface label, while the interface section edge is acquired by the next one. D. Cozzolino et al. [24] proposed a new technique for the detection of image splicing using a feature-based algorithm. In this technique, the co-occurrence of images is used to compute local features. Those local features are then used to extract feature parameters. Since spliced and host images can exhibit different properties, the expectation–maximization algorithm, together with the segmentation, is used for learning purposes.

In view of the above studies, most of the techniques used for forgery detection are based on handcrafted methods for feature extraction, which are highly dependent on the individual undertaking the task. The development of deep learning-based methods has led to automatic feature extraction. The use of deep learning thus removes possible human errors and increases the efficiency and reduces the time complexity of the model.

3. Proposed System Architecture

In this study, we proposed a deep learning-based approach for the identification of forged images. We proposed an architecture using ResNet50v2 as our base model, and we used the YOLO CNN weights for transfer learning. This approach enabled us to train the model with meaningful weights. We used pre-trained weights of the YOLO CNN object detection model to initialize our ResNet50v2-based proposed architecture, which saved a considerable amount of training costs, as we initialized our model with meaningful pre-trained weights.

Figure 3 presents the basic architecture of ResNet50v2, in which initially batch normalization is performed, followed by an activation function and the weights being updated. Then we performed the batch normalization, ReLU activation function. After the activation function, the weights were optimized. The basic difference from the ResNet50v2 architecture is that we used pre-activation of the weight layers instead of post-activation. ResNet50v2 was developed in such a way that it removes the nonlinearity, hence clearing a path from the input to the output as a means of an identity connection. Version 2 of the ResNet module also applies the batch normalization and the activation function before the weights are multiplied. The overall proposed system is shown in Figure 4.

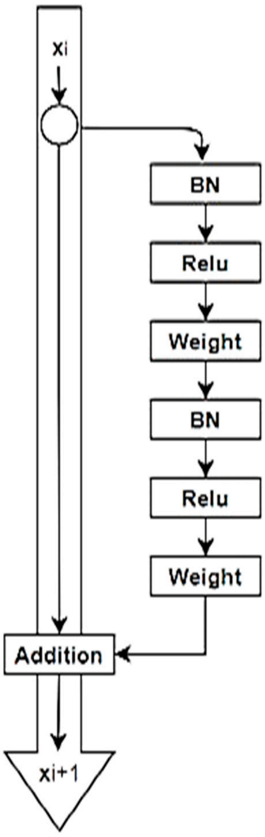

Figure 3. ResNet50v2 architecture.

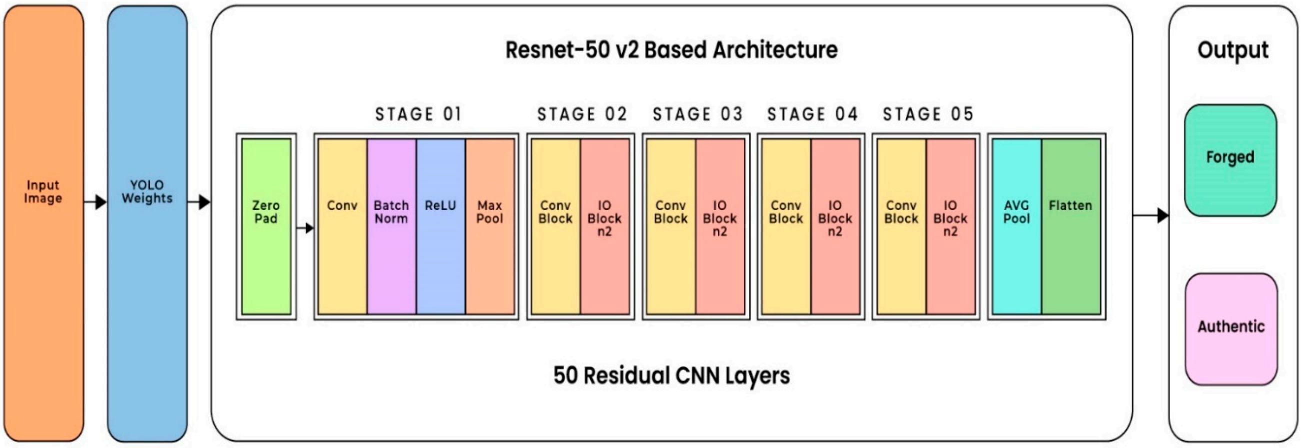

Figure 4. ResNet50v2-based architecture for proposed system.

Since the input and output dimensions are not the same, the residual block function is defined in Equation (1), as follows:

$$y = F(x,\ \{W_i\}) + W_s x \tag{1}$$

In Equation (1), F is the residual block function, x represents the input image, and W represents the pre-trained weights of the YOLO CNN. Since the residual block function does block mapping with zero extra paddings, changing the dimensions, Resnet produces significantly better results.

We also addressed the degradation problem by the utilization of a deep residual learning framework. The desired underlined mapping is fitted into the stack layers. Formally, if we represent underlined mapping as $H(x)$, then the mapping of another nonlinear stacked layer is $F(x) := H(x) - x$. We determined that optimizing the original is more difficult compared to residual mapping. It is easy to push zero residuals compared to a nonlinear layer stack. The formulation of $F(x) - x$ is considered a shortcut connection. It refers to skipping one or more layers [25]. If identity mappings are generated from added layers, the information will be able to flow through the network, allowing any layer to serve as an original input and reducing training error.

Transfer Learning

Transfer learning is considered an application of deep learning that enabled us to incorporate the pre-trained weights of an existing model on large data containing hundreds, if not thousands, of classes. Since these weights are pre-trained for large and challenging datasets, a high-end computing machine with GPU is required for this purpose, and it can take days or weeks to train and validate the model. The transfer learning approach reduces the cost of training a model from scratch and allows for achieving more accurate results in less time. As stated in [26], transfer learning is an approach by which we could optimize our model, which prevented us from training the model from scratch, and hence, improved the performance. Figure 5 shows the comparison using transfer learning in a CNN versus no transfer learning. From this graph, we can see that at the start, the transfer learning-based deep learning model performed better compared to training a model from the scratch. In this paper, we present a deep learning-based architecture that uses a transfer learning technique to utilize the weights of a YOLO CNN.

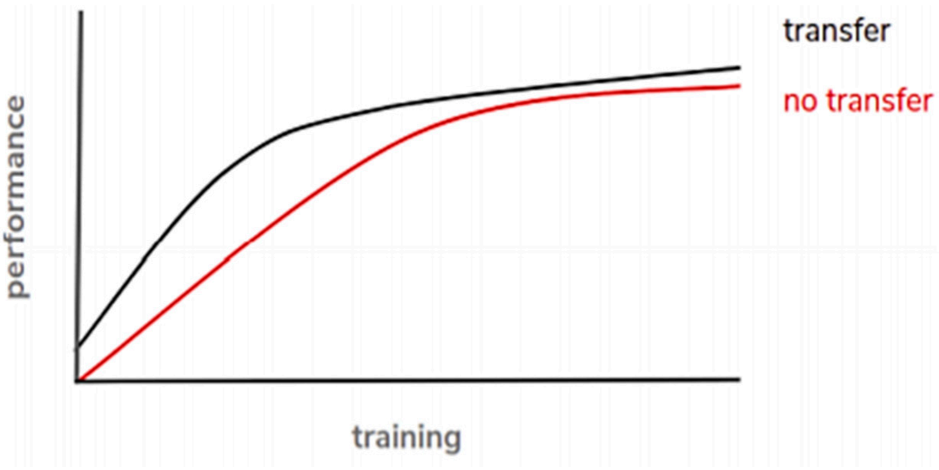

Figure 5. Transfer learning vs. no transfer learning [26].

In our proposed system, ResNet50v2 is used as a basic convolution model, which comprises five stages. With separate convolution and identity blocks, each block consists of three convolution layers, and the identity block also has three convolution layers. There are over 23 million parameters that can be trained for the ResNet50v2 model.

We used the CASIA ITDE v1 and v2 datasets for this purpose, which consisted of two classes of original and forged images. The dataset was divided into training and testing sets.

Figure 4 shows the deep learning-based proposed architecture of our proposed system. The proposed model takes an input image and uses pre-trained YOLO CNN weights to detect the authenticity of an image.

4. Dataset

There are some software applications available for the detection of tempered images such as Adobe Photoshop. Since a public standard dataset was not available before CASIA datasets, researchers worked and experimented with their proposed approaches on limited examples [27–29]. Since there was no benchmark dataset, it was very difficult to compare the accuracy and effectiveness of a technique. Now, some benchmark datasets are available on forgery detection, such as CASIA_v1 and CASIA_v2. In this study, we performed our experiments and analysis on two benchmark datasets, CASIA_v1 and CASIA_v2 [9].

4.1. Preparation of Dataset

Dong et al. [9] collected a dataset and named it the CASIA Image Tampering Evaluation Database. Adobe Photoshop CS3 version 10.0.1 is used to generate all the color images for the tampered database. In addition, this dataset contains CASIA_v1 and CASIA_v2 for the image tampering detection evaluation database. CASIA_v1 contains 1721 color images, and CASIA_v2 contains 12,323 color images. CASIA_v1 only focuses on splicing as a tampering technique, and hence, all the tampered images in this dataset are classified as spliced tampered images. The size of images in the CASIA_v1 database is fixed as 384 × 256 and they are stored in JPEG format. Apart from that, tampered images in CASIA_v2 are more comprehensive compared to CASIA_v1. We discuss the construction of CASIA_v1 in detail in the following section.

4.2. CASIA ITDE v1

The CASIA ITDE v1 dataset is a collection of 1721 color images that are 384*256 pixels in size. The images are in JPEG format, as shown in Table 1. These images are further divided into two sets, as follows:

1. Forged set;
2. Authentic set.

Table 1. Difference between CASIA_v1 and CASIA_v2 datasets.

Dataset	Distribution	Size	Format
CASIA_v1	46–54% distribution between authentic and forged images	374 × 256 color image	JPEG
CASIA_v2	55–45% distribution between authentic and forged images	from 320 × 240 to 800 × 600 color image	TIFF, JPET, BMP,
Columbia (Uncompressed)	50% distribution each for authentic and forged images	1152 × 768 color images	TIFF
Columbia (Compressed)	52–48% distribution between authentic and forged images	128 × 128 color images	BMP

After the division of the dataset into two subsets, the forged set contained 921 images, whereas there were 800 images in the authentic dataset, so the authentic set contained 46% and the rest belonged to the forged set. We used two sources to generate the authentic dataset. Most of the images were taken from the Corel image dataset [30]. The Corel database is a well-known image database used for the development of many professional applications. Based on the image content, the authentic set contains images of eight types (scene, texture, nature, plant, article, character, animal, and architecture). Since the generation of forged images requires the modification of original images, forged images also contain the eight types mentioned above. The crop and paste tool in Adobe Photoshop was used for the generation of forged sets from authentic images. Figures 6 and 7 show some examples taken from the CASIA_v1 dataset. After generating the spliced image, it is stored using the same filename. Table 2 shows the statistical features of the spliced images.

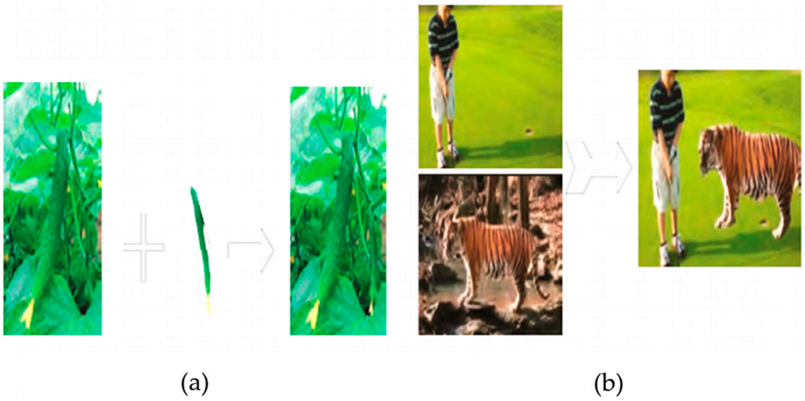

Figure 6. (**a**) Forged Image in CASIA ITDE v1 (**b**) Forged Image in CASIA ITDE v1.

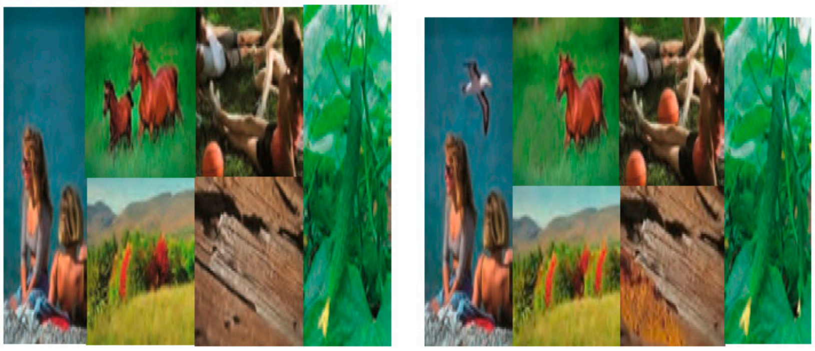

Figure 7. Forged image in CASIA ITDE v1.

Table 2. Statistical information of spliced images.

Category		Count
JPEG Format		921
Source of Tampered Region(s)	Different images	470
	Same image	451
Preprocessing Manipulations	Resize	206
	Rotation	25
	Resize and distortion	27
	Rotation and distortion	3
	Distortion	53
	Rotation and resize	45
	Rotation, distortion, and resize	0
Manipulations without pre-processing		562
Tampered Region Boundaries	Rectangular	169
	Circular	114
	Arbitrary	536
	Triangular	102

Spliced images are generated based on the following criteria:

1. Spliced image regions are either generated from the same authentic image or a combination of different authentic images.
2. Spliced region shapes can be changed and customized using the Adobe Photoshop palette.
3. Rotation, scaling, and other operations can be applied to cropped images before being added to spliced images.
4. Spliced regions are generated with different spliced region sizes.
5. The authentic set also contains texture images since forgery can easily be noticeable with text. Thus, we cropped a random region for texture images.

4.3. CASIA ITDE v2

The CASIA_v2 dataset is an extended version of the CASIA_v1 dataset. This dataset contains 12,323 samples and these samples are divided into two subsets.

The tampered set contains 5123 images, whereas the authentic set has 7200 images. The CASIA_v2 dataset is more comprehensive compared to the CASIA_v1 dataset. This dataset contains images with different dimensions, ranging from 320 × 240 to 800 × 600, and contains uncompressed images, which include TIFF and BMP samples. The authentic subset is constructed from the dataset proposed by Corel et al. [30] and a tampered subset was generated after blurring the authentic subset.

The spliced region's edge or any other region can be used with the blurring technique for the generation of tampered images. This is the unique difference between the CASIA_v1 and CASIA_v2 tampered sets. Figures 8 and 9 show examples of tampered images in the CASIA_v2 dataset [9].

The following rules were considered while generating tampered images to make this dataset more comprehensive:

1. Photoshop was used to define realistic images as close to human vision as possible.
2. Tampered images were generated either from two different authentic images or from the same authentic image.
3. Cropped images were further processed with distortion, rotation, and scaling before being inserted to generate a realistic image.

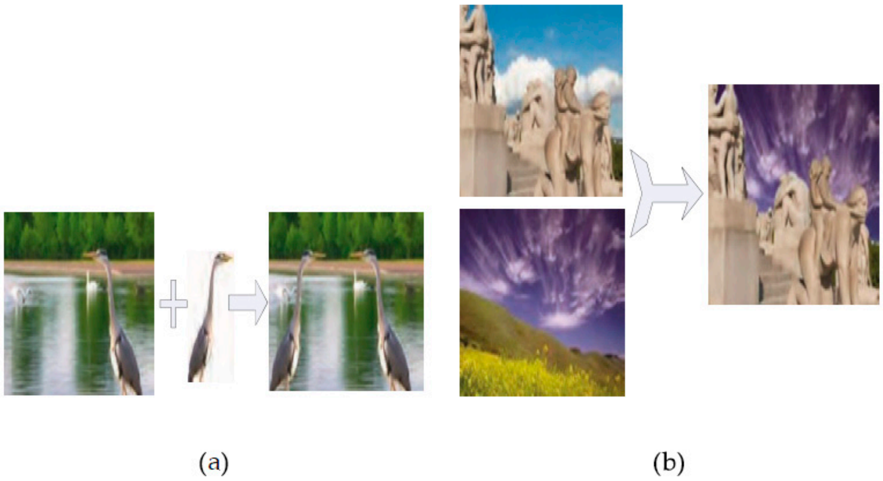

Figure 8. (**a**) Tampered image in CASIA_v2 dataset (**b**) Tampered image in CASIA_v2 dataset.

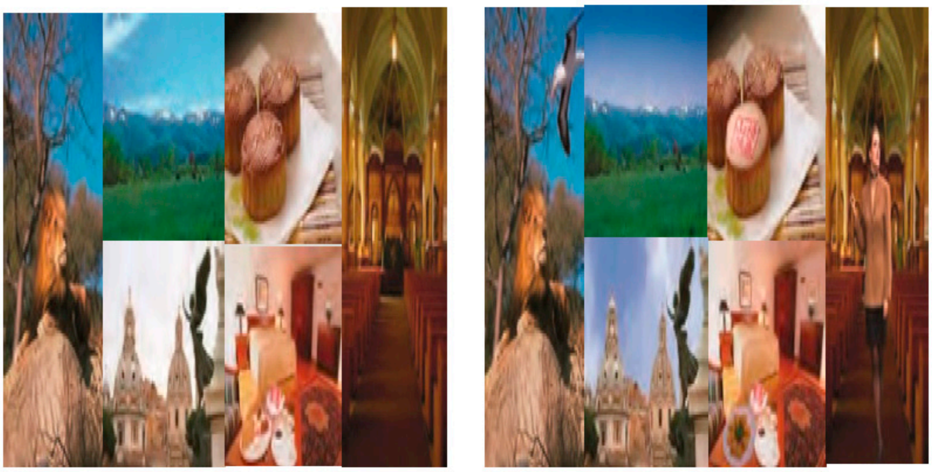

Figure 9. Comparison of authentic and tampered images.

While generating forged images, different size images are generated for tampered sets. Table 1 shows the detailed differences between the CASIA_v1 and CASIA_v2 datasets.

4.4. Dataset Evaluation

A test was designed to evaluate the quality of the generated tampered dataset. Thirty (30) people were given 100 images to identify each image as tampered with or not using the naked human eye. They correctly identified the tempered images with an average accuracy of 59%, which illustrates that these tampered images are more realistic compared to the Colombia uncompressed images dataset. Table 1 shows the comparison of the CASIA_v1, CASIA_v2, and Columbia (compressed and uncompressed) datasets.

5. Experimental Results and Discussion

In order to evaluate the proposed system, we conducted different experiments to demonstrate and evaluate the effectiveness of the proposed deep learning-based approach. We also compared our proposed technique with several generic tampering detection techniques with different publicly available datasets.

We performed all the experiments on the publicly available CASIA_v1 and CASIA_v2 datasets. Table 3 shows the system specifications of the hardware that was used for conducting the experiments. Each dataset contains copy/move and splicing images. CASIA_v2 is a more comprehensive and challenging dataset compared to the CASIA_v1 dataset because it contains images of different sizes and formats.

Table 3. System specification details.

Name	Details
Operating system	Windows 10
Programming language	Python
Model	Asus ROG 702 VM
RAM	32 GB
Cores	16

5.1. System Specification

We used the Windows 10 operating system and ASUS ROG 702 VM for conducting the experiments. All the system details are provided briefly in Table 3.

5.2. Results Discussion and Comparison

We compared the convergence performance of the proposed deep learning-based approach with and without transfer learning. Figure 10 shows the comparison of results of the CASIA_v1 and CASIA_v2 datasets with transfer learning. The blue color represents the accuracy per epoch for the CASIA_v1 dataset, whereas the orange color represents the accuracy for the CASIA_v2 dataset. A total of 25 epochs were executed to check the effectiveness of our proposed technique. After the execution of the first epoch, we saw a strong positive trend for both datasets. We achieved an accuracy of 99.3% with CASIA_v2, and 81% with the CASIA_v1 dataset. Since CASIA_v2 is considered to be a more comprehensive and benchmark dataset, the experimental results obtained with the CASIA_v2 dataset are good compared to those from the CASIA_v1 dataset. The equation for calculating the accuracy is given in Equation (2).

$$Accuracy = \frac{TP + TN}{Total\ Samples} \tag{2}$$

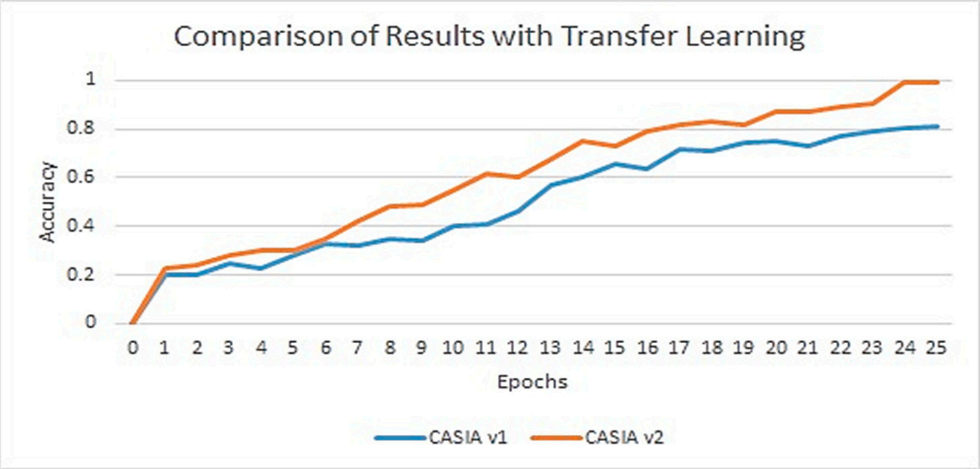

Figure 10. Testing process visualization with transfer learning.

The comparison of the results of our proposed architecture without using transfer learning is shown in Figure 11. This approach had adverse effects on our proposed system architecture, as the model learned from scratch with randomly initialized weights. Using random initialization for weights tends to increase the complexity, cost, and training time. The random initialization is non-meaningful, so the model had to do a lot of work to update the weights. On the other hand, the use of transfer learning enabled us to use pre-trained weights that are meaningful. The model did not have to be trained from scratch, thus decreasing the training time, cost, complexity, and increasing the accuracy. In our proposed deep learning-based architecture, we used the pre-trained YOLO CNN weights. You Only Look Once (YOLO) is a real-time, state-of-the-art object detection system that is trained on thousands of different objects. Figures 10 and 11 explain the performance evaluation of the proposed technique with and without transfer learning, respectively. We achieved an accuracy of 80% without the use of transfer learning in the CASIA_v2 dataset, while we obtained an accuracy of 69.3% for the CASIA_v1 dataset.

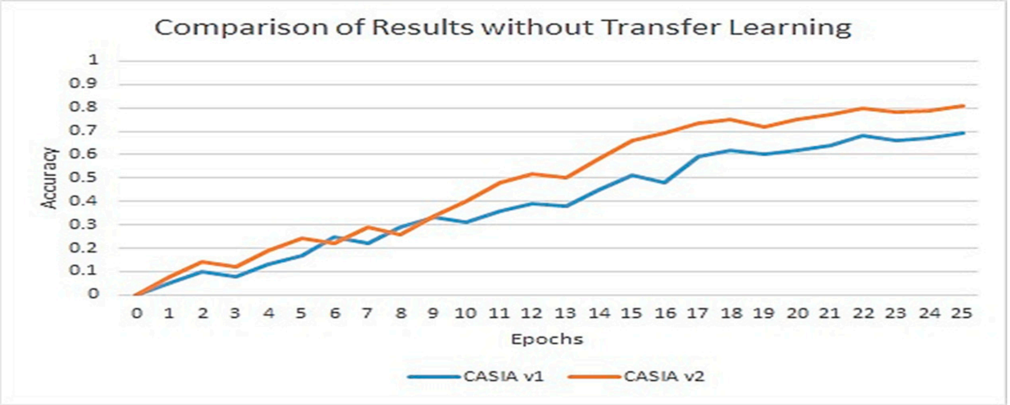

Figure 11. Testing process visualization without transfer learning.

Figures 12 and 13 present the comparison of the losses between transfer learning and without transfer learning among the two datasets. An increase in loss is seen when we did not use the pre-trained weights of the YOLO CNN model. The loss was slightly higher in the CASIA_v1 dataset compared to CASIA_v2. Furthermore, we compared our proposed method with existing architectures, which can be seen in Table 4, which presents the type of tampering targeted, along with the methodology used to detect the tampering. We also list the advantages and disadvantages of the respective architectures and their obtained accuracies.

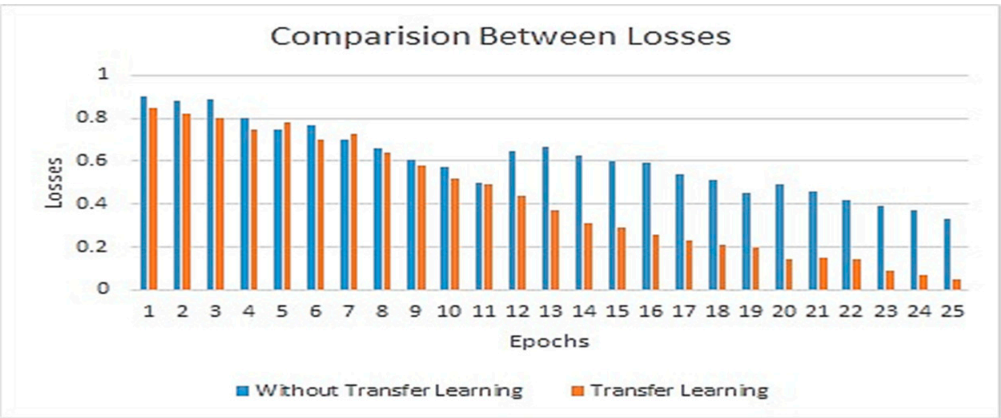

Figure 12. Comparison between losses on CASIA_v2 dataset.

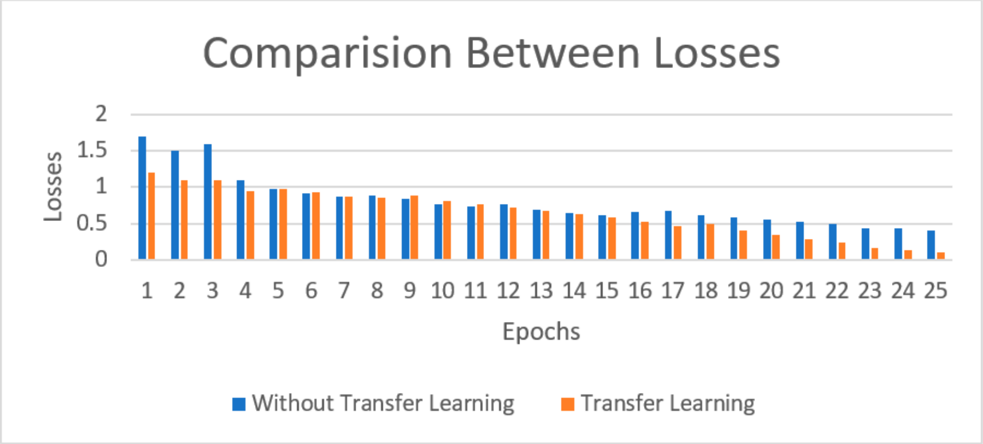

Figure 13. Comparison between losses with CASIA_v1 dataset.

Table 4. Evaluation and characteristics of the proposed model.

Folds	Loss	Accuracy (%)
Fold 1	0.08	99.15
Fold 2	0.09	99.31
Fold 3	0.05	99.45
Fold 4	0.09	99.33
Fold 5	0.07	99.45
Average	0.076	99.33

Cross-validation is a technique used to divide the data into a given number of sets and train the model on each set. Table 4 shows the evaluations and performance of five folds. The average accuracy obtained was 99.33% and the average loss was 0.076. According to the comparative analysis with the methodologies mentioned in Table 5, the proposed architecture performed better compared to the traditional tampering detection techniques.

Table 5. Comparison between existing approaches.

Study	Tampering Targeted	Methodology	Dataset	Advantages/Disadvantages	Accuracy
[19]	Cut/paste, copy/move	CNN	CASIA_v1, CASIA_v2, and Columbia DVMM	Advantage: The compressed feature of the test set is mined. A feature fusion technique is also used to attain good results. Disadvantage: Model complexity.	98.04%
[31]	Cut/paste, copy/move	Mask R-CNN, ResNet-101	Columbia, Cover	Advantage: Better performance compared to other techniques. Disadvantage: Unable to follow contours.	93% precision for Cover dataset and 97% precision for Columbia dataset

Table 5. *Cont.*

Study	Tampering Targeted	Methodology	Dataset	Advantages/Disadvantages	Accuracy
[23]	Image splicing	MFCN, edge probability map, and surface probability	CASIA_v1	Advantage: The proposed methodology performs better than current splicing. Disadvantage: Uses the training set for image assessment on new images.	0.52 MCC score
[32]	Cut/paste	CNN	Dresden database	Advantage: The proposed methodology uses a CNN for mining features using camera point hints. Disadvantage: Not able to identify localization and camera model traces.	Localization Accuracy is 81% and Detection Accuracy is 82%
[33]	Cut/paste, JPEG double compression	Multi-domain CNN and RGB features of DCT	UCID	Advantage: The proposed methodology uses a CNN for localizing and categorizing patches of images that are compressed. Disadvantage: It does not make use of CNNs to perceive various types of compressions.	95%
[34]	Cut/paste	Autoencoder and landscapes with noise	Images taken from 7 electronic devices	Advantage: The proposed methodology uses obtains fair results. Disadvantage: It does not explore the use of several degrees of freedom.	0.41 F-Measure
[20]	Cut/paste	RRU-Net and Image residuals	Columbia, CASIA	Advantage: Attained somewhat good results on tamper detection without any preprocessing. Disadvantage: Latent discriminative features are not expressed.	93.94%
[35]	Copy/move, cut/paste	SAE and Daubechies wavelet decomposition	CASIA_v1, CASIA_v2 and Columbia	Advantage: Attained somewhat good results on tamper detection without any preprocessing. Disadvantage: The areas that need to be recognized must have to be manually inferred. There is not precise detection of the areas that have been inferred.	90.09%

Table 5. *Cont.*

Study	Tampering Targeted	Methodology	Dataset	Advantages/Disadvantages	Accuracy
[22]	Attacks that are a combination of transformations	AlexNet Model, CNN	MICC-F220	Advantage: The proposed model uses SVM as a classifier to attain good accuracy. Disadvantage: Less suitable for other datasets.	93.94%
[18]	Median filtering, AWGN, Gaussian blurring	CNN and error filter predictions	Images taken from 12 unique cameras	Advantage: The proposed model uses a CNN for manipulation detection and attains good accuracy.	99.10%
[36]	Cut/paste and median filtering	CNN and Median filtering residuals	Boss base, UCID, Dresden, BOSS RAW, NRCS Gallery	Advantage: The proposed model gives considerably good results. Disadvantage: Less suitable for other datasets.	85.14%
Our Proposed Work	Image splicing	ResNet50v2 and YOLO weights	CASIA_v1, CASIA_v2	Advantage: Reduces training time and uses ResNet-based architecture.	99.3%

5.3. Future Work

Forgery detection is an ever-growing problem that needs constant improvement of the mechanisms used to detect tampered images. There are multiple techniques used for creating forged images, such as copy/paste, lighting conditions, image splicing, and retouching. In this research, we are focusing on the detection of the images that have been spliced. In the future, our proposed technique can be extended to the detection of multiple types of forged images and can be tested on multiple datasets as well. Furthermore, this work can be taken forward specifically to improve the effectiveness of multiple types of forgery detection by a single model.

6. Conclusions

Image forgery detection is a very challenging problem. In this era of technological advancement, we need to be able to distinguish between real and tampered images. In this study, we proposed a deep learning-based approach for image forgery detection. The proposed model is based on ResNet50v2 architecture, which uses residual layers; thus, using this architecture increases the detection rate of tampered images. Using this approach also provides the benefit of transfer learning by using the pre-trained weights of the YOLO CNN model. The use of transfer learning enabled us to train our model more efficiently, as we initialized our proposed model by meaningful assigning weights. This reduced the training time and complexity of the model and makes the architecture more efficient. We evaluated our proposed architecture on benchmark datasets, CASIA_v1 and CASIA_v2. We also compared the performance of our system with and without the use of transfer learning. We obtained an accuracy of 99.30% with the CASIA_v2 dataset for the forgery detection problem. The results of the comparison with the existing methods show the superiority of the proposed system. The proposed system will help in the image manipulation detection domain and also paves the way for future research in detecting multiple types of image forgery manipulations.

Author Contributions: Conceptualization, E.U.H.Q. and T.Z.; data curation, E.U.H.Q., T.Z., and A.A.; formal analysis, E.U.H.Q., T.Z., and A.A.; funding acquisition, A.A.; methodology. E.U.H.Q.; project administration, T.Z. and A.A.; resources, A.A.; software, E.U.H.Q.; supervision, T.Z. and A.A.; validation, E.U.H.Q., T.Z., and A.A.; visualization, E.U.H.Q.; writing—original draft, E.U.H.Q.; writing—review and editing. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by Security Research Center at Naif Arab University for Security Sciences (Project No. SRC-PR2-01).

Acknowledgments: The author would like to express their deep thanks to the Vice Presidency for Scientific Research at Naif Arab University for Security Sciences for their kind encouragement of this work.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Fridrich, J.; Soukal, D.; Lukás, J. Detection of Copy-Move Forgery in Digital Images. *Int. J. Comput. Sci.* **2003**, *3*, 55–61.
2. Yerushalmy, I.; Hel-Or, H. Digital Image Forgery Detection Based on Lens and Sensor Aberration. *Int. J. Comput. Vis.* **2011**, *92*, 71–91. [CrossRef]
3. Dirik, A.E.; Memon, N. Image tamper detection based on demosaicing artifacts. In Proceedings of the 2009 16th IEEE International Conference on Image Processing (ICIP), Cairo, Egypt, 7–10 November 2009; pp. 1497–1500. [CrossRef]
4. Christlein, V.; Riess, C.; Jordan, J.; Riess, C.; Angelopoulou, E. An Evaluation of Popular Copy-Move Forgery Detection Approaches. *IEEE Trans. Inf. Forensics Secur.* **2012**, *7*, 1841–1854. [CrossRef]
5. Zampoglou, M.; Papadopoulos, S.; Kompatsiaris, Y. Large-scale evaluation of splicing localization algorithms for web images. *Multimed. Tools Appl.* **2016**, *76*, 4801–4834. [CrossRef]
6. Varga, D. Multi-Pooled Inception Features for No-Reference Image Quality Assessment. *Appl. Sci.* **2020**, *10*, 2186. [CrossRef]
7. Kawahara, J.; Bentaieb, A.; Hamarneh, G. Deep features to classify skin lesions. In Proceedings of the 2016 IEEE 13th International Symposium on Biomedical Imaging (ISBI), Prague, Czech Republic, 13–16 April 2016; pp. 1397–1400.
8. Bai, X.; Yang, M.; Huang, T.; Dou, Z.; Yu, R.; Xu, Y. Deep-Person: Learning discriminative deep features for person Re-Identification. *Pattern Recognit.* **2020**, *98*, 107036. [CrossRef]
9. Dong, J.; Wang, W.; Tan, T. CASIA Image Tampering Detection Evaluation Database. In Proceedings of the 2013 IEEE China Summit and International Conference on Signal and Information Processing, Beijing, China, 6–10 July 2013.
10. Mahdian, B.; Saic, S. A bibliography on blind methods for identifying image forgery. *Signal Process. Image Commun.* **2010**, *25*, 389–399. [CrossRef]
11. Farid, H. Image forgery detection. *IEEE Signal Process. Mag.* **2009**, *26*, 16–25. [CrossRef]
12. Lanh, T.V.; Chong, K.; Emmanuel, S.; Kankanhalli, M.S. A Survey on Digital Camera Image Forensic Methods. In Proceedings of the 2007 IEEE International Conference on Multimedia and Expo, Beijing, China, 2–5 July 2007; pp. 16–19. [CrossRef]
13. Barad, Z.; Goswami, M. Image Forgery Detection using Deep Learning: A Survey. In Proceedings of the 2020 6th International Conference on Advanced Computing and Communication Systems (ICACCS), Coimbatore, India, 6–7 March 2020; pp. 571–576.
14. Wu, Y.; Abd-Almageed, W.; Natarajan, P. BusterNet: Detecting Copy-Move Image Forgery with Source/Target Localization. In *Lecture Notes in Computer Science*; Springer: Berlin, Germany, 2018; pp. 170–186.
15. Manjunatha, S.; Patil, M.M. Deep learning-based Technique for Image Tamper Detection. In Proceedings of the 2021 Third International Conference on Intelligent Communication Technologies and Virtual Mobile Networks (ICICV), Tirunelveli, India, 4–6 February 2021; pp. 1278–1285. [CrossRef]
16. Schaefer, G.; Stich, M. UCID: An uncompressed color image database. *Storage Retr. Methods Appl. Multimed.* **2003**, *5307*, 472–480.
17. Amerini, I.; Ballan, L.; Caldelli, R.; DEL Bimbo, A.; Serra, G. A SIFT-Based Forensic Method for Copy–Move Attack Detection and Transformation Recovery. *IEEE Trans. Inf. Forensics Secur.* **2011**, *6*, 1099–1110. [CrossRef]
18. Bayar, B.; Stamm, M.C. 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 2016; pp. 5–10.
19. Rao, Y.; Ni, J. A deep learning approach to detection of splicing and copy-move forgeries in images. In Proceedings of the 2016 IEEE International Workshop on Information Forensics and Security (WIFS), Abu Dhabi, United Arab Emirates, 4–7 December 2016; pp. 1–6. [CrossRef]
20. Bi, X.; Wei, Y.; Xiao, B.; Li, W. 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 2019; pp. 30–39.
21. Zhan, Y.; Chen, Y.; Zhang, Q.; Kang, X. 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 2017. [CrossRef]
22. Doegar, M.D.A.; Gaurav, K. CNN Based Image Forgery Detection Using Pre-trained AlexNet Model. *Int. J. Comput. Intell. IoT* **2019**, *2*, 6.

23. Salloum, R.; Ren, Y.; Kuo, C.-C.J. Image Splicing Localization using a Multi-task Fully Convolutional Network (MFCN). *J. Vis. Commun. Image Represent.* **2018**, *51*, 201–209. [CrossRef]
24. Cozzolino, D.; Poggi, G.; Verdoliva, L. 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 2015; pp. 1–6.
25. Bishop, C.M. *Neural Networks for Pattern Recognition*; Oxford University Press: Oxford, UK, 1995.
26. Lopes, R.d.l.F. *Wild Data Part 3: Transfer Learning*; Stratio Big Data Inc.: Madrid, Spain, 2018.
27. Bayram, S.; Sencar, H.T.; Memon, N. An efficient and robust method for detecting copy-move forgery. In Proceedings of the 2009 IEEE International Conference on Acoustics, Speech and Signal Processing, Taipei, Taiwan, 19–24 April 2009; pp. 1053–1056. [CrossRef]
28. Qu, Z.; Qiu, G.; Huang, J. Detect Digital Image Splicing with Visual Cues. In Proceedings of the Information Hiding: 11th International Workshop, IH 2009, Darmstadt, Germany, 8–10 June 2009; Revised Selected Papers. Springer: Berlin, Germany, 2009; pp. 247–261.
29. Wei, W.; Dong, J.; Tan, T. Effective image splicing detection based on image chroma. In Proceedings of the 2009 16th IEEE International Conference on Image Processing (ICIP), Cairo, Egypt, 7–10 November 2009; pp. 1257–1260. [CrossRef]
30. Corel Database. Available online: http://www.coreldraw.com/ (accessed on 4 August 2021).
31. Wang, X.; Wang, H.; Niu, S.; Zhang, J. Detection and localization of image forgeries using improved mask regional convolutional neural network. *Math. Biosci. Eng.* **2019**, *16*, 4581–4593. [CrossRef] [PubMed]
32. Bondi, L.; Lameri, S.; Guera, D.; Bestagini, P.; Delp, E.; Tubaro, S. Tampering Detection and Localization through Clustering of Camera-Based CNN Features. In Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), Honolulu, HI, USA, 21–26 July 2017; pp. 1855–1864.
33. Amerini, I.; Uricchio, T.; Ballan, L.; Caldelli, R. Localization of JPEG Double Compression through Multi-domain Convolutional Neural Networks. In Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), Honolulu, HI, USA, 21–26 July 2017; pp. 1865–1871.
34. Cozzolino, D.; Verdoliva, L. Single-image splicing localization through autoencoder-based anomaly detection. In Proceedings of the 2016 IEEE International Workshop on Information Forensics and Security (WIFS), Abu Dhabi, United Arab Emirates, 4–7 December 2016; pp. 1–6.
35. Zhang, Y.; Goh, J.; Win, L.L.; Thing, V. Image Region Forgery Detection: A Deep Learning Approach. In Proceedings of the Singapore Cyber-Security Conference (SG-CRC), Singapore, 14–15 January 2016; IOS Press: Singapore, 2016.
36. Chen, J.; Kang, X.; Liu, Y.; Wang, Z.J. Median Filtering Forensics Based on Convolutional Neural Networks. *IEEE Signal Process. Lett.* **2015**, *22*, 1849–1853. [CrossRef]