# On the benchmarking of ResNet forgery image model using different datasets


Faria Hossain
*School of Engineering*
*London South Bank University*
London, United Kingdom
hossaif5@lsbu.ac.uk

Tasos Dagiuklas
*School of Engineering*
*London South Bank University*
London, United Kingdom
tdagiuklas@lsbu.ac.uk



***Abstract*—This paper presents the benchmarking and improvement of the ResNet image forgery model using three different datasets (CASIA, Columbia, and LSBU). The model is based on classification, where forgery images have been edited using cut-paste modification technique.The images are categorized to check if the algorithm can successfully identify the difference between the original and the forgery image. All images have been pre-processed with Gray-Edge detectors to obtain get better classification results. Experimental results have shown that the Gray-edge technique has improved the accuracy across all image datasets.**



***Index Terms*—Image Forgery, ResNet, LSBU-Columbia-CASIA image datasets**


## I. Introduction

Image forging refers to copying and pasting an image's content into another image [1]. This procedure become more frequent. Low-cost and high-resolution cameras have made it possible for everyone to take and save digital photos. In addition, even a non-expert may change a picture using simple photo editing software. Enhancing, retouching, splicing, morphing, copy-move, cut-paste, and erase-fill techniques are all key features of image forgery approaches.

Different image editing tools could be used in scientific, industrial, and military sectors where image modification is essential [2]. Such tools could be used to enhance a picture, such as by adding saturation, blur, tone, brightness, and so on.There is no negative impact on the image's meaning from these improvements. Retouching is the most popular method of preparing an image for a final performance [3]. It's mostly utilized to remove or intensify aspects of a picture to draw the reader's attention. The procedure is often difficult and requires certain abilities. A good retouching approach may completely change the image from top to bottom [4].The process of combining multiple pictures into a single view is known as image splicing. A single picture is created from a group of images at the final. Image transformation is related to image height and width. The surgery is sometimes referred to as a "morph" for short. An image-morphing computer approach gradually transforms images into new ones. Cut-Paste and Copy-Move belong to Splicing, and Erase-Fill belongs to Inpainting [4]. Editing any image using cut-paste means that some part of another image is taken and posted on the old image [1]. The new image is a forgery image. Such type of image editing may benefit people in different businesses but editing an image is sometimes used for unethical purposes as well [1]. A human sometimes may detect a forgery image very quickly but detecting so many images at a time would be difficult [1]. Similarly, sometimes machines can better predict the editing than human beings [1].

This is due to the fact that an image can be utilized as legal evidence in forensics investigations and numerous other fields. The goal of pixel-based image forgery detection is to check the authenticity of digital images without knowing anything about the original [1]. The impact of forgery can be disturbing in terms of law and order, and culprits can use this to save themselves, and innocent people could get hurt by the wrong forgery evidence images [1].Image Modification is a Frequently Asked Question in Several Disciplines. Forgeries are famous in medical science, media, sports, criminal investigation, image forensics, and other vital industries in which picture authenticity is crucial. Mainly, methods based on deep learning may produce media content nearly undetectable to human eyes and very close to being realistic. Image manipulation has interested businesses, social media platforms, and other media industries [5]. Standards for media content and metadata are studied on JPEG have set up a WG on Fake Media to define annotations of media changes [6]. It covers the key subjects and case studies where learning and assessment are significant. Deep Learning algorithms like ResNet, BusterNet, CNN, and other classification algorithms have widely been used to detect the Forgery image [4]. ResNet with 50 layers has been widely used in the past to detect cut-paste forgery images [4]. This paper proposes benchmarking improvement of a ResNet method to detect improvement in the detection of cut-paste models. Such an improvement is based on edge detection. Benchmarking has been carried out on three different image datasets with standards. The structure of the paper is the following: Section II is discussed the literature review of the work. The project methodology is provided in part III. Datasets overview is presented in section IV while section V discusses the complete outcome of the project. Finally, in Section VI, the author acknowledges and cites all of the sources used in the paper's research.

## II. LITERATURE REVIEW

### A. Cut-Paste Forgery Images

Cut-paste refers to the joining of either two or more images together. There is an original image from which the part that has been changed has been cut out, swapped with the piece from a different reference image, and added back to the original image to make the tempered image. Cut-paste tampering indicators are often less visible than copy-move region duplication clues [1]. Contrast enhancement on the original two images could be required by a picture counterfeiter if they had been taken in illumination variations than the final material picture.

### B. ResNet Image Classification

The algorithm is discussed in detail. The ResNet and its variations have accomplished noteworthy triumphs in different PC visions. Deep Learning algorithms like ResNet, BusterNet, CNN, and other classification algorithms have widely been used to detect the Forgery image [8]. ResNet with 50 layers has been commonly used to see cut-paste forgery images [2]. The researchers likewise present an examination of CIFAR-10 with 100 and 1000 layers. The profundity of portrayals is of focal significance for some visual acknowledgment undertakings. Deep nets are groundworks of their entries to ILSVRC COCO 2015 competitions. Despite its outcome in making slope courses through building blocks, the data correspondence of halfway layers of blocks is overlooked. To resolve this, the authors [9], propose to present a controller module as a memory component to separate correlative highlights of the middle layers, which are additionally taken care of by the ResNet. Specifically, the controller module uses of convolutional RNNs (e.g., LSTMs or GRUs), which are demonstrated to be great at extricating spatial-fleeting data. The new managed network is dubbed ResNet. The controller module can be effectively executed and annexed to any ResNet design. Exploratory outcomes on three picture arrangement datasets have exhibited the promising execution of the proposed engineering contrasted and the standard ResNet, crush and excitation ResNet, and other best-in-class structures.

To guarantee harmonious intermingling, preparing a CNN without any preparation is extremely demanding for more computational resources. Tweaking a CNN that has been pre-prepared utilizing, for example, a gigantic arrangement of marked clinical datasets is a reasonable other option. In a paper, a similar study was finished using pre-prepared models, for instance, VGG-19 and ResNet-50, as against preparing without any preparation [8]. To reduce overfitting, information expansion and dropout regularization were utilized. With a review of 92.03%, authors have shown that the pre-prepared models with legitimate fine-tuning are equivalent to Iyke-Net, a CNN prepared from scratch [8].

## III. METHODOLOGY

### A. ResNet Model

The model comprises 5 phases, each with a convolution and Personality block. Every convolution block has three convolution layers, and every identity block has three convolution layers. The ResNet-50 has north of 23 million teachable boundaries [9]. As of now, networks are turning out to be increasingly mind-boggling, from a few layers to many layers. The principal benefit of neural networks is that they can communicate extremely complex capabilities. It can gain highlights from various degrees of reflection, for example, edge highlights at lower levels and complex elements at higher levels. However, the utilization of networks isn't generally successful because there is an exceptionally enormous obstruction - the vanishing of slopes: in incredibly profound organizations, angle signals will often move toward zero rapidly, making the angle plunge process very slow. Specifically, during the time spent inclination drop, the weight framework item activity should be done in each step of the back spread from the last layer to the top layer so that the slope will drop dramatically to 0. Subsequently, during the time spent preparing, it will be tracked down that with the increment of the number of layers, the pace of inclination declines increases. Therefore, by extending the organization, even though it can communicate any intricate capability, with the increment of organization layers [9].

Neural Networks like the well-known ResNet-50 model, a convolutional CNN that is 50 layers profound. ResNet [10] is an ANN that stacks remaining blocks on top of one another to frame an organization. As an immediate consequence of these headways, it has become workable for PC vision models to outperform people in effectively tackling various issues connected with picture acknowledgment, object location, face declaration, picture characterization [11] in such a manner, the presentation of profound convolutional brain organizations or CNNs merits extraordinary notice. These organizations have been widely utilized for dissecting visual symbolism with exceptional precision. Be that as it may, while it provides us with the choice of adding more layers to the CNNs to tackle more confounded errands in PC vision, it accompanies its arrangement of issues. It has been seen that preparing the brain networks turns out to be more troublesome with the expansion in the quantity of added layers, and now and again, the exactness lessons also [12]. It is here that the utilization of ResNet accepts significance. More profound networks are more challenging to prepare. With ResNet, it becomes conceivable to outperform the hardships of preparing deep networks. ResNet has numerous variations that sudden spike in demand for a similar idea; however, have various quantities of layers. Resnet50 indicates the variation that can work with 50 network layers. There are skip associations in ResNet. Skip associations work in two ways. They, first and foremost, mitigate the issue of evaporating inclination by setting up a substitute easy route for the slope to go through [13]. Also, they empower the model to get familiar with a character's capability. This guarantees that the higher layers of the model play out no more terribly than the lower layers. The remaining blocks make it significantly more straightforward for the layers to learn personality capabilities. Thus, ResNet works on the proficiency of profound networks with additional layers while limiting the level of mistakes. As such, the skip associations add the results from past layers to the effects of stacked layers, making it conceivable to prepare a lot of different organizations than beforehand conceivably [9].

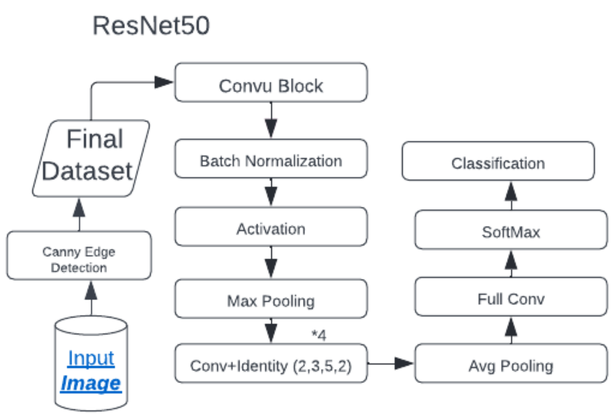


Fig. 1. Project Methodology [14]

### B. Gray-Edge Detection

We did not use the exact images in models. Before moving towards benchmarking, we have applied the edge detection method on images. The following steps are completed. All Pictures are refreshed through the foundation disposal strategy. The edge discovery strategy is utilized to choose the pertinent picture and eliminate all the other things. Watchful edge identification technique from cv2(OpenCV) is applied to recognize the edges [15] impeccably. This procedure impeccably distinguishes every one of the pictures.

1. Apply a Gaussian channel to smooth the picture to eliminate the commotion [16]

2. Find the force inclinations of the picture

3. Apply inclination greatness thresh holding or lower bound slice off concealment to dispose of fake reaction to edge discovery

4. Apply a twofold limit to decide on expected edges

5. Track edge by hysteresis: Conclude the recognition of edges by smothering the wide range of various edges that are powerless and not associated with areas of strength for with [17].

This Gray edge detection is looped through all the images in Train and Test. The process is repeated for every image that is being used in every ResNet method that we are going to discuss further.

D. Accuracy Calculations The accuracy is calculated by checking the correct predictions with total test data. The formula for the accuracy is stated below.

$$Accuracy = CorrectPredictions/TrueData \quad (1)$$

## IV. Dataset

Three datasets are used in this project. The basic explanation of each dataset is provided below:

### A. LSBU

This dataset consists of 1000 original and 3000 forgery images generated from authentic images. The original photos have been retrieved from publicly available repositories, and LSBU dataset considers both high and low resolutions. The forgery images have been created using three different methods: cut-paste, erase-filling, and copy-move. To create the fake images, both pre-processing and post-processing have been utilized. This includes sharpening, enhancing color and size, blurring, and adjusting exposure. Resizing, rotation, sampling, and sharpening are all included in the next generation of fake photos. It is available on the IEEE data port [18].

### B. CASIA

CASIA's team established a web portal in 2009 in response to the increased need for larger assessment datasets and more realistic altered photos [19]. The CASIA 1.0 collection contains 1721 photos; 800 are authentic, and 921 have been tempered. JPEG and TIFF file formats are used to store compressed and uncompressed pictures in the CASIA 2.0 dataset, which contains images ranging from 240×160 to 900×600 pixels in size. Total color photographs are 12,614 and include 7491 originals and 5123 retouched shots. Tampered photos have been post-processed to improve their influence on the second dataset, CASIA 2.0.

### C. Columbia

It was created in 2004 when the Columbia picture splicing detection validation dataset was made public [20]. This is the first publicly accessible dataset for studying image manipulation. Grayscale photographs are only included in this. Nine hundred thirty-three original blocks and 912 cut-paste images were created in Adobe Photoshop. They are all of the same size (128×128) and are saved as BMP files [20]. A new dataset was created in 2006 to solve the shortcomings of the previous one. This updated dataset includes 183 color photographs and 180 cut-and-paste graphics. Uncompressed TIFFs of the color photographs are provided in sizes ranging from 757×568 to 1152×768. The cut-paste tampering method was the focus of the two Columbia datasets. The dataset is used in this project accordingly.

The dataset is divided so that 216 images from Columbia dataset are used in training. There are 48 images used in the test dataset. Three hundred ninety-six images are used for training the CASIA dataset, and 54 images are used in the test dataset. In the case of the LSBU dataset,144 images are used in the training dataset, and 44 images are used in the test dataset.

TABLE I
TRAIN TEST SPLIT OF THE DATASETS

| **Dataset** | **Image Technique** | | | |
|---|---|---|---|---|
| | ***Forgery Train*** | ***Original Train*** | ***Forgery Test*** | ***Original Test*** |
| LSBU | 72 | 72 | 22 | 22 |
| CASIA | 198 | 198 | 27 | 27 |
| Columbia | 108 | 108 | 24 | 24 |

[a]Dataset Counts.

## V. RESULTS AND DISCUSSIONS

Fig. 2. Original Image

Fig. 3. Canny-Edge detection for Original Image

Figure 2 presents the original image from the dataset, and Figure 3 is the edge detected image. In Figure 3 it can be seen that the points where the image has edges are detected clearly.

Fig. 4. Forgery Image

Fig. 5. Canny-Edge detection for Forgery Image

Figure 4 presents the Forgery image of Image 2 from the dataset, and Figure 5 is the edge detected image. In Figure 5 it can be seen that the points where the image is edited are detected clearly, and it will be easy for the algorithm to detect this image as a forgery image. The cap of the cartoon is edited and ResNet algorithm may now distinguish clearly.

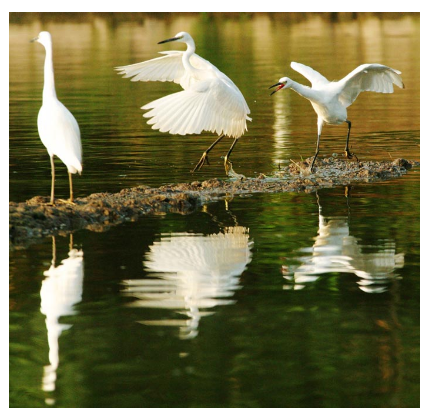

Fig. 6. Original Image

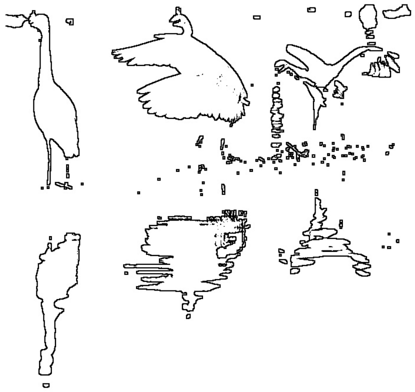

Fig. 7. Canny-Edge detection for Original Image

Figure 6 presents the original image from the dataset, and Figure 7 is the edge detected image. In Figure 7 it can be seen that the points where the image has edges are detected clearly. All of the birds and their shadow is carefully detected.

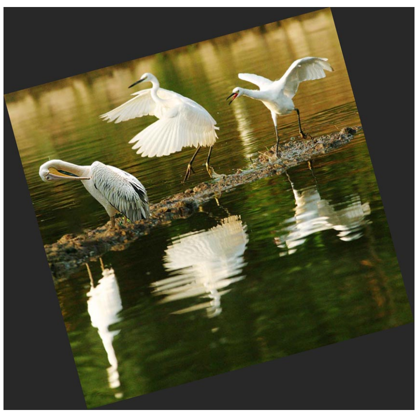

Fig. 8. Forgery Image

Fig. 9. Canny-Edge detection for Forgery Image

Figure 8 presents the Forgery image of Image 6 from the dataset, and Figure 9 is the edge detected image. In Figure 9, it can be seen that the points where the image is edited are detected clearly, and it will be easy for the algorithm to detect this image as a forgery image. The birds in the Forgery images are a bit flipped, and the canny edge has accurately detected the birds' position.

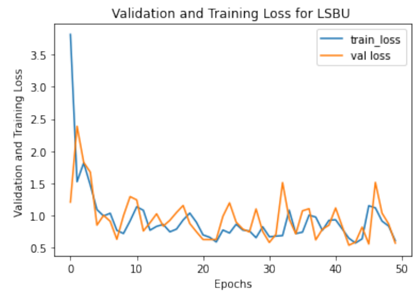


Fig. 10. Validation and Training Loss for LSBU

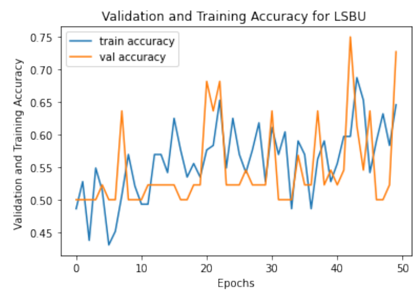


Fig. 11. Validation and Training Accuracy for LSBU

Figure 10 presents the train and validation loss for the LSBU dataset and the ResNet algorithm. Figure 11 presents the accuracy changes through the epochs. In Figures 10 and 11, the accuracy and loss are presented. The algorithm has detected 75% of the predictions correctly.

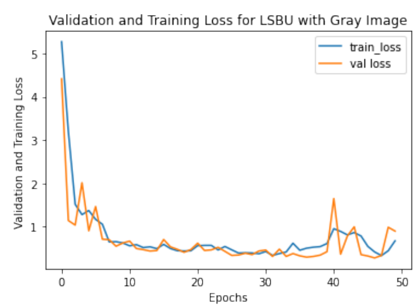


Fig. 12. Validation and Training Loss for LSBU with Canny Edge

Figure 12 presents the train and validation loss for the Canny Edge LSBU dataset with and the ResNet algorithm. Figure 13 presents the accuracy changes through the epochs. In Figures 10 and 11, the accuracy and loss are presented. The algorithm has detected 97% of the predictions correctly. Hence, The Canny-Edge detection has improved the accuracy.

Figure 14 presents the train and validation loss for the CASIA dataset and the ResNet algorithm. Figure 15 presents the accuracy changes through the epochs. In Figures 14 and

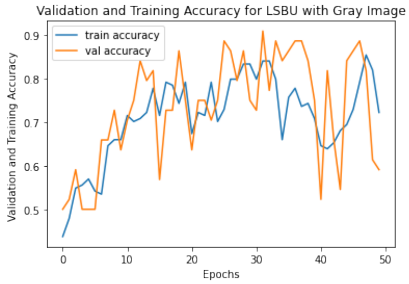


Fig. 13. Validation and Training Accuracy for LSBU with Canny Edge

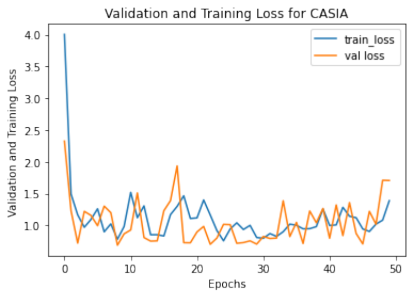


Fig. 14. Validation and Loss Accuracy for CASIA

15, the accuracy and loss are presented. The algorithm has detected only 57% of the predictions correctly. This is a very low value for such an important dataset.

Figure 16 presents the train and validation loss for the Canny Edge CASIA dataset with and the ResNet algorithm. Figure 17 presents the accuracy changes through the epochs. In Figures 16 and 17, the accuracy and loss are presented. The algorithm has detected 65% of the predictions correctly. Hence, The Canny-Edge detection has improved the accuracy a lot. It is at-least better than 57% of without Canny-Edge.

Figure 18 presents the train and validation loss for the Columbia dataset and the ResNet algorithm. Figure 19 presents the accuracy changes through the epochs. In Figures 18 and 19, the accuracy and loss are presented. The algorithm has detected 78% of the predictions correctly. This accuracy is the best accuracy for any data set without Gray-Edge detection.

Figure 20 presents the train and validation loss for the Canny Edge Columbia dataset with and the ResNet algorithm.

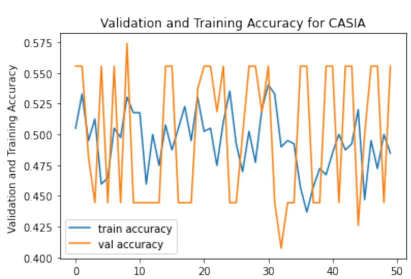


Fig. 15. Validation and Training Accuracy for CASIA

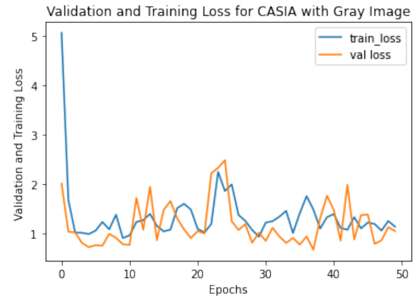


Fig. 16. Validation and Training Loss for CASIA with Canny Edge

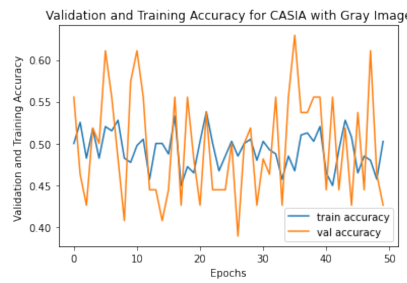


Fig. 17. Validation and Training Accuracy for CASIA with Canny Edge

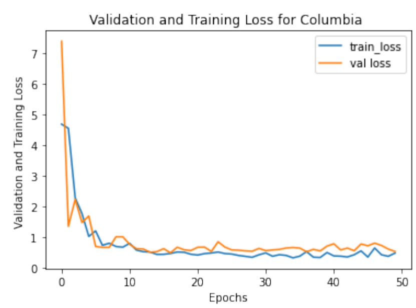


Fig. 18. Validation and Training Loss for Columbia

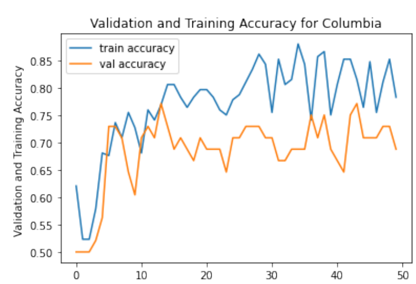


Fig. 19. Validation and Training Accuracy for Columbia

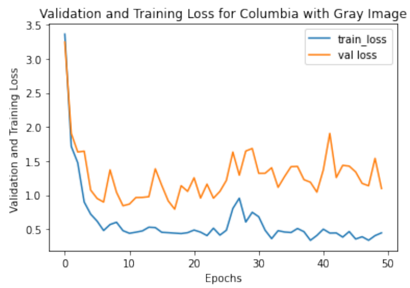


Fig. 20. Validation and Training Loss for Columbia with Canny Edge

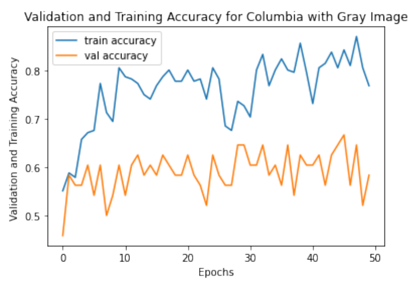


Fig. 21. Validation and Training Accuracy for Columbia with Canny Edge

Figure 21 presents the accuracy changes through the epochs. In Figures 20 and 21, the accuracy and loss are presented. The algorithm has detected 65% of the predictions correctly. Hence, The Canny-Edge detection has not improved the accuracy. Rather it becomes lower than the original one. It shows that the Canny Edge is best for LSBU and CASIA but simple images are best for Columbia dataset.

TABLE II
COMBINED ACCURACY USING RESNET 50 (PERCENT)

| Dataset | Image Techniques | |
|---|---|---|
| | ***Without Canny Edge*** | ***Using Canny Edge*** |
| LSBU | 75.1 | 97.0 |
| CASIA | 57.2 | 65.40 |
| Columbia | 78.1 | 65.5 |

[a]ResNet50.

Table 2 has the accuracy details for all three datasets. In the case of not using gray-edge, there was much less accuracy for LSBU and CASIA. In the future, more data can be added so that proper verification of LSBU data classification can be done without stating the low accuracy case.

## VI. CONCLUSION

This paper presents the benchmarking of ResNet to deal with cut-paste using three different image datasets. Among these datasets, LSBU shows high accuracy. It is due to the implementation of the Background elimination method.The Gray-Canny Edge detection technique is applied to all images. Our model with Canny Edge using ResNet on Forgery images has proved much better than simple ResNet on the same data. The accuracy of the LSBU test data increased up-to the 97% from 75% using the canny edge detection technique. The Gray-Canny Edge detection is best for LSBU due to its low error rate, which signifies the ability to identify just existent edges.

## REFERENCES


[1] Lilei Zheng, Ying Zhang, Vrizlynn L.L. Thing, A survey on image tampering and its detection in real-world photos, Journal of Visual Communication and Image Representation, Volume 58, 2019, Pages 380-399, ISSN 1047-3203, https://doi.org/10.1016/j.jvcir.2018.12.022. (https://www.sciencedirect.com/science/article/pii/S104732031830350X)

[2] Edit ¿ Erase fill (no date) Autodesk.com. Available at: https://knowledge.autodesk.com/search-result/caas/CloudHelp/cloudhelp/2019/ENU/MSHMXR/files/GUIDA39BD725-0326-480B-AA34-4387964AB846-htm.html (Accessed: 09 May 2022).

[3] Brown, L. (2021) High end retouching techniques for beginners, Pixels NYC. Available at: https://pixelsnyc.com/high-end-retouching-techniques/ (Accessed: 09 May 2022).

[4] JK. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition," CoRR, vol. abs/1512.03385, 2015, [Online]. Available: http://arxiv.org/abs/1512.03385

[5] A. P. Jagtap and H. A. Hingoliwala, "Advanced Techniques for Image Forgery Detection," Int. J. Comput. Appl., vol. 146, no. 10, pp. 20–25, Jul. 2016, doi: 10.5120/ijca2016910936.

[6] JPEG - JPEG Fake Media (no date) Jpeg.org. Available at: https://jpeg.org/jpegfakemedia/index.html (Accessed: 26 May 2022).

[7] Temmermans, F., Bhowmik, D., Pereira, F. and Ebrahimi, T., 2021, August. JPEG Fake Media: a provenance-based sustainable approach to secure and trustworthy media annotation. In Applications of Digital Image Processing XLIV (Vol. 11842, p. 118420L). International Society for Optics and Photonics. 117

[8] J. Xu, Y. Pan, X. Pan, S. Hoi, Z. Yi, and Z. Xu, "RegNet: Self-Regulated Network for Image Classification.," IEEE Trans. Neural Netw. Learn. Syst., vol. PP, Mar. 2022, doi: 10.1109/TNNLS.2022.3158966.

[9] Y. Liu, H. Fan, F. Ni, and J. Xiang, "ClsGAN: Selective attribute editing model based on classification adversarial network," Neural Netw., vol. 133, pp. 220–228, 2021.

[10] A. V. Ikechukwu, S. Murali, R. Deepu, and R. C. Shivamurthy, "ResNet-50 vs. VGG-19 vs training from scratch: A comparative analysis of the segmentation and classification of Pneumonia from chest X-ray images," Glob. Transit. Proc., vol. 2, no. 2, pp. 375–381, 2021, doi: https://doi.org/10.1016/j.gltp.2021.08.027.

[11] X. Jin, Y. Zou, and Z. Huang, "An imbalanced image classification method for the cell cycle phase," Information, vol. 12, no. 6, p. 249, 2021.

[12] B. Li and Y. He, "An improved ResNet based on the adjustable shortcut connections," IEEE Access, vol. 6, pp. 18967–18974, 2018.

[13] Y. Jiang, L. Chen, H. Zhang, and X. Xiao, "Breast cancer histopathological image classification using convolutional neural networks with small SE-ResNet module," PloS One, vol. 14, no. 3, p. e0214587, 2019.

[14] S. Showkat and S. Qureshi, "Efficacy of Transfer Learning-based ResNet models in Chest X-ray image classification for detecting COVID-19 Pneumonia," Chemom. Intell. Lab. Syst., vol. 224, p. 104534, 2022.

[15] Matsuyama, "A Deep Learning Interpretable Model for Novel Coronavirus Disease (COVID-19) Screening with Chest CT Images," J. Biomed. Sci. Eng., vol. 13, pp. 140–152, Jan. 2020, doi: 10.4236/jbise.2020.137014.

[16] E. Matsuyama, "A Deep Learning Interpretable Model for Novel Coronavirus Disease (COVID-19) Screening with Chest CT Images," J. Biomed. Sci. Eng., vol. 13, pp. 140–152, Jan. 2020, doi: 10.4236/jbise.2020.137014.

[17] S. Malik and T. Kumar, "Comparative Analysis of Edge Detection between Gray Scale and Color Image," Commun. Appl. Electron., vol. 5, pp. 38–43, May 2016, doi: 10.5120/cae2016652230.

[18] Hossain, F. (2022) "Forgery Image Dataset." IEEE Dataport.

[19] O. M. Al-Qershi and B. E. Khoo, "Evaluation of Copy-Move Forgery Detection: Datasets and Evaluation Metrics," Multimed. Tools Appl, vol. 77, no. 24, pp. 31807–31833, Dec. 2018, doi: 10.1007/s11042-018-6201-4.

[20] Columbia Image Splicing Detection Evaluation Dataset , "https://www.ee.columbia.edu/ln/dvmm/downloads/AuthSplicedDataSet /AuthSplicedDataSet.htm"