Fake colorized and morphed image detection using convolutional neural network

Neetu Pillai*
Assistant Professor, Department of Computer Engineering, PHCET, Maharashtra, India

Received: 19-January-2020; Revised: 26-February-2020; Accepted: 28-February-2020
©2020 Neetu Pillai. This is an open access article distributed under the Creative Commons Attribution (CC BY) License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Abstract
As the technology develops, utilization of the phony pictures is at highest, So, as per an overview most of the pictures stored on the server or in the cloud are being transformed or counterfeit. As a result, it is difficult to identify whether the images stored are real or not. So not many systems are existing today which are equipped to tell whether pictures are fake or not. Earlier Histogram based and feature extraction-based methods were used to identify fake images. The neural network is being the most advanced technology distinguishes the phony pictures by studying various features of the image and learning the procedure of faking an image. So as a minor advance towards this, proposed system uses features like dark channel, bright channel, RGB channel and alpha channel. Using Gaussian distribution, we can anlayze the edges of the images to identify phony images. Deep layer analysis is performed using convolutional neural system along with the fuzzy classification procedure to improve the likelihood of recognizing phony pictures. Through this research it is concluded that convolutional neural networks (CNN) based fake colorized image detection showcases a better result than histogram based and feature extraction based fake colorized image detection.

Keywords
Convolution neural network, Gaussian distribution, Fuzzy classification.

1. Introduction
A normal smartphone on a normal client produces approximately 1000 photos. As the quantity of pictures continues expanding, there is additionally an inundation of a practically equivalent number of phony pictures. In the present generation, images can be modified using advanced editing software technologies. The morphed images are planned with evil purposes to ruin the first proprietor by people with malevolent expectations. [1]

Because of countless produced pictures, it is important to have the option to produce an original picture from the manufactured ones. This has driven numerous information researchers to examine in this field and that has empowered a far more noteworthy comprehension of the frauds that are submitted. The commonly using image fabrication method is an Image splicing and CopyMove technique. The copy move strategy is increasingly regular in contrast to the Image splicing method. With the new image editing applications like Photoshop, Photo Editor, the procedure of faking an image has been improved.

Individuals with such malevolent aims would use that cash for progressive accursed purposes. Hence, finding a fake and morphed image is necessary to maintain authenticity of the original image. Previously histogram based and feature extraction-based methods were used to identify fake images. Through research it is concluded that using modern technology like Convolutional Neural Network, have enhanced the performance of the fake image detection at higher rates.

AI is designed according to the human instinct; it can give significant knowledge into the information a lot quicker and human-like. To imitate human knowledge the calculation needs to make a judgment through instinct, which must be conceivable through experience. To this end, AI utilizes a lot of methods, for example, artificial neural networks, and instant based learning. These apparatuses help the AI calculation build up a kind of human-like instinct by gathering information and building a semantic connection between its parts [1].

Neural systems are called all things considered in light of the fact that these systems are designed
according to the human mind and its operations. This is finished by imitating how the mind chips away at the fundamental level. The neuron is the most essential unit of calculation in the mind. It is the littlest unit of the mind that is equipped for processing at an inborn level. A neuron is a cell equipped for transmitting an electrical drive on excitation.

Colorization of grayscale pictures is anything but a perplexing undertaking for the inventive human cerebrum. The significant level insight indispensable for this procedure is actually why the test of building up a completely programmed colorization calculation is such a huge errand. The system of programs colorization changes two indiscrrept perceptions into a plan reasoning. Initially, semantic data matters. So as to colorize irregular pictures, a framework must see the semantic setup of the scene notwithstanding limiting articles. Profound convolutional neural networks (CNN) can help as instruments to solidify confinement and semantic parsing into a colorization framework.

The broad utilization of cell phones and cameras has prompted a plenitude of pictures which thus has expanded the likelihood of people with fiendish expectations to alter and produce pictures by leaving next to no or no follow. The scientists have done a broad examination of the picture falsification procedures and their discoveries and give us a comprehensive assessment of the location techniques used in deciding the kind of imitation, duplicate move, re-testing, grafting, modifying and so forth [1].

The main objective of this research paper is to enhance fake image detection process using modern technology like CNN along with fuzzy classification with accuracy and with less period of time.

2. Literature review

Hsu et al. [2] proposed a deep learning-based approach to identify fake images generated using generative adversarial networks (GAN) through contrastive loss. The author has proposed a fake feature network, which is trained using the pairwise learning for differentiating the features of real and fake images and outputs the probability whether the image is real or fake. The proposed system performs better than state-of-the-art fake image detectors.

Research gap: The proposed system fails when the fake features of the image of a new generator are different from that in the training phase. At such times, both the fake image and image detector needs to be re-trained. Next limitation of proposed technology is a collection of datasets.

Yan et al. [3] introduces a system that detects recolored image using the deep discriminative model. The system takes the original image and along with two derived inputs based on illumination consistency and inter-channel correlation of the original input and outputs the probability whether it is required or not. The deep discriminative model is used to enhance the performance of the system with greater accuracy. The author proposes an end-to-end deep discriminative neural network to distinguish natural images from recolored images, which captures more comprehensive features.

Research Gap – The system detects only recolored images and designing of effective network architecture is required for searching high-level cues for better distinguishing.

Li et al. [4] developed a system for detecting fake colorized images by leveraging neural network. The author analyzes statistical difference between natural images and corresponding fake images. The author has also proposed a simple feature extraction technique to measure similarity between the normalized histogram distribution of both images. The author has conducted experiments by training the dataset which is created by three state of the art colorization technique to enhance the performance and robustness of the system.

Guo et al. [5] presented the various sorts of phonies that should be possible with a picture and the different ways that can be utilized to distinguish the fakes. Be that as it may, this doesn't consider the colorization of grayscale pictures. This is hard to recognize as there aren't numerous systems that have been created for the identification of phony colorization of the pictures. In this way, to give an answer for this issue, the creators proposed two distinct techniques for the location of the phony colorized pictures, the first is the fake colorized image detection-histogram (FCID-HIST), and the other one is the FCID-feature encoding (FE) based fake colorized image detection. These systems have been demonstrated to be superior to the customary discovery techniques by a major edge.

Research Gap – As with any learning technique, a large portion of the identification relies upon the sort of preparing given to the framework and is restricted.
from that point of view, as it can't identify falsification in any event, when the picture is manufactured, however the framework is constrained to its comprehension.

Gragnaniello et al. [6] disclosed the expansion of web-based life as it transforms into a very groundbreaking specialized apparatus. Because of the blast of internet-based life, there has been a tremendous flood of controlled pictures. This has prompted a scourge and there are scientists that have been attempting to battle this issue with the consideration of CNN's in their indicators. As an ongoing investigation have demonstrated that CNN based recognition techniques are powerless against assaults. In this way, to improve this impact the scientists have broken down the antagonistic assaults done on different diverse CNN based location frameworks and tried their exhibition satisfactorily.

Research Gap – the investigation focuses just on the assaults that occur in a CNN based system, which is extremely restricted and have not proposed any medicinal frameworks that can be utilized without a CNN based methodology.

Zhao et al. [7] communicates worries over the creation of information as loads of information is being produced each day. Huge data is a promising field that can procedure these gigantic measures of information to get gainful bits of knowledge from it. Be that as it may, has been transcendentally damaged with the presence of a ton of phony, controlled or produced pictures. Consequently, the creators have endeavored to isolate the phony pictures and distinguish the first picture from a specific dataset. The method being utilized by the analysts depends on Convolutional Neural Networks as it is very powerful in the region of picture preparing. The system has been tried widely and brought about the affirmation that this strategy id one the most proficient strategies.

Research Gap –The profound learning techniques that are being utilized resemble whatever other profound learning strategy whose presentation prevalently relies on how viably it has been prepared and that relies upon the ability of the designer.

Bachtiar et al. [8] states that because of the expansion in web-based life and the measure of picture control strategies, there has been an enormous scale inundation of phony and controlled pictures. This likewise brings about a great deal of phony tricks being spread around, making frenzy and mania.

While there has a great deal of calculations that have been grown as of late with the end goal of identification of the imitations, yet they have missed the mark regarding a wide range of situations. The analysts present a system of recognition of phony pictures with the assistance of the amalgamation of the calculation and web advances. This procedure has been demonstrated to be powerful and profoundly productive than its partners.

Research Gap – The analysts have utilized just the de-mosaicing strategy for the recognition of the phony pictures that are being produced, however haven't used it with any of the AI procedures to improve the framework much further.

Wang et al. [9] presented a multi-watermarking is an idea that different watermarks are implanted inside a similar division of the host spread subtly, which can be utilized to secure the responsibility for follow merchants or synergistic writers in media content conveyance. However, there are different open issues in multi-watermarking, the method to distinguish watermarks, the limitations to approve numerous, watermarks' exhibitions, and so on.

As per multi-watermarking models, two kinds of multi-watermarking half breed decoders, i.e., locally ideal and ideal, are proposed, and the models with Gaussian and Laplacian dispersions are given. Moreover, the normal BER, as the proportion of the most invaluable decoders in model two and model one, is examined hypothetically.

In this paper, two sorts of multi-watermarking models which are half and half and multiplicative are displayed for conveying multi-bit data, named model one and model two. The half breed multiplicative implanting rule of model one fits in the application where the watermark vitality is dissipated in general unique spread, while that of model two for the application where the watermark vitality centers around neighborhood districts of the first spread.

Yan et al. [10] described about movement obscure which is regularly caused because of camera movement, which influences the vast majority of the PC vision assignments. So as to deblur scenes including normal, face, content, and low-light pictures dark channel prior (DCP) is utilized. DCP has certain downsides and doesn't bolster bits estimation when overwhelming pixels which are brilliant rules input picture. Free picture bright pixels are less inclined to be brilliant, considerably after the
deblur procedure. From these perception scientists have delineated this wonder numerically and have later characterized it as the bright channel prior (BCP). This exploration exhibits a method for expelling the haze from the pictures which additionally praises the traditional procedures by giving a reasonable picture. The method abuses the two kinds of pictures that is, pictures that are both splendid channel and dim channel earlier. It makes powerful rebuilding efforts by using both the channels and makes a superior deblurred picture. Here in this paper outrageous channel earlier is considered to deblur the picture which frequently neglects to obscure the picture for the dim foundation. The blurring will distinguish the edges of the phony hued pixels. In our proposed model Gaussian distribution is utilized for this reason. To battle the issue of splendid pixels lessening in their splendor after the deblurring procedure, the creators plan a procedure by a blend of this earlier with the DCP. This earlier is named ECP or successful earlier for the recuperation of inactive pictures for piece estimate. The outcomes are true to form and the calculation beats the traditional calculations and can be utilized related to them.

Gadhiya et al. [11] clarifies that the utilization of computerized pictures in the domain of restorative diagnostics has been an extraordinary shelter to the general public and the therapeutic field as there can be a great deal of sicknesses and issues that can be identified now without the utilization of any intrusive methodology. This was beyond the realm of imagination prior as the vast majority of the issues required broad medical procedure and biopsies to determine certain infections. There have likewise been instances of altering and the genuineness of the restorative picture has come into question. To enhance this impact the scientists have exhibited a calculation for the recognition and restriction of the altering in the pictures with the assistance of discrete wavelet change. The strategy is effective and exceptionally in its tasks.

Research Gap – This procedure has been tried on different databases that have been available with the controlled pictures. Yet, there are extremely less medicinal picture datasets that are accessible that can be utilized to test this calculation of its maximum capacity.

Larsson et al. [12] describes about Colorization of grayscale pictures is anything but a perplexing errand for the innovative human mind. An individual is just required to review that grass is green and the sky is blue, for different articles, the psyche is allowed to envision a few potential hues. The elevated level comprehension indispensable for this procedure is actually why the test of building up a completely programmed colorization calculation is such a gigantic errand. The procedure of programs colorization changes two rash perceptions into a plan theory. Right off the bat, semantic data matters. So as to colorize arbitrary pictures, a framework must see the semantic design of the scene notwithstanding restricting articles. Profound CNNs can help as instruments to unite confinement and semantic parsing into a colorization framework. The article shows a framework that exhibits the ordinary capacity to consequently shading grayscale pictures. Two imaginative contributions empower this advancement: a shading histogram expectation structure that oversees ambiguities and vulnerability normal for colorization while forestalling restorative ancient rarities, and a profound neural design that is prepared to coordinate semantically huge highlights of various unpredictability into colorization.

Research Gap – the creators have intended to make the procedure significantly increasingly instinctive and help cause the procedure to show signs of improvement by improving the learning situations.

Huang et al. [13] describes JPEG, which is the most frequently used image formats in our daily life. JPEG images consist of RDH technique which has series of application and has greater significance. If any modification is done in JPEG images it introduces distortion compared to uncompressed image. Increasing the storage size is disproportional to the length of embedded message as a result the file size of the marker image may increase significantly. The paper presents a new block selection strategy which results in good visual quality and less storage size of the JPEG file. Other RDH schemes can be utilized to improve their performance. The proposed technique has strong practicability because high embedding capacity and good visual quality can be easily obtained. Meanwhile, the storage size of the original image file can be well preserved.

Research Gap – the researchers have planned to use block detection technology in their future endeavors to increase the system's overall accuracy.

Li et al. [14] introduces a technique for the picture with CMF comprises of no, not exactly a few areas which are undefined. CMF possibly directed by a
falsifier seeking after, either to improve the spatial visualization of the picture or to cover reality. Ordinary individuals may disregard this fraud activity when the counterfeiter purposefully shrouds the altering. As a picture is once in a while fashioned without a target, the duplicate move locales ought to have a specific significance. In this light, the scientists propose to section the test picture into various non-covered patches. Thus, the CMFD can be executed by concurring these patches, as long as the replicating source locales and the sticking objective doesn't exist in a similar fix. This paper essentially centers around the location of the section duplicated from one picture into another picture to counterfeit a picture. Here a source picture is expected to distinguish the fabrication. This paper proposed a CMFD procedure built up in picture division. In spite of the fact that the CMF areas are recognized essentially by approximating the key focuses inferred in the picture while arranging the proposed plan as a keypoint-based one is beyond the realm of imagination. It very well may be viewed as a blend of both existing plans on the grounds that in the two phases of the coordinating procedure both key focuses and pixel highlights are utilized.

Research Gap—The proposed technique is computationally burdening and should be possible in a superior manner with the joining of parallel processing.

Qureshi and Deriche [15] investigated the coming of computerized pictures and their multiplication in this new web age. The broad utilization of cell phones and cameras has prompted a plenitude of pictures which thusly has expanded the likelihood of people with underhanded aims to alter and manufacture pictures by leaving almost no or no follow. The analysts have done a broad examination of the picture falsification procedures and their recognitions and furnish us with a comprehensive assessment of the location techniques used in deciding the kind of fabrication, duplicate move, resampling, grafting, modifying and so forth.

The scientists in the wake of having done a broad assessment of the systems have gone over a perception that the analysts who have been effectively creating techniques and calculations for altering recognition have not been focusing on the exploration that goes into counter-crime scene investigation or hostile to legal sciences. This region of research centers around the unequivocal evacuation of altering follows is looked for after by these falsification recognition techniques.

3. Proposed methodology

Below diagram depicts working of the proposed system for enhancing the fake image detection (Figure 1).

![Diagram](image)

**Figure 1** Proposed system working diagram

**Step 1: Feature Extraction**

In this step, base features of images are extracted like dark channel, bright channel, alpha channel and RGB Channel. For extraction of these features the system considers one input and traverse through each pixel to get a signed integer (SIGN). Right shifting the SIGN integer by 24 bits and performing AND operation with a hexadecimal value results to an
Alpha channel value, same way right shifting SIGN integer by 16,8,0 bits and performing AND operation with a hexadecimal value results to Red, Green and Blue channel value.

**Step 2: Edge Estimation**
Through edges of an image we can determine whether the image is fake or real. If the image is real then edges of the image will be intact, whereas if the image is fake, then the values of the edge image will be fluctuating. For identifying edges of the image edge detection algorithm is used along with ROI (Region of Interest) Identification and ROI extraction.

ROI edges of an image are identified using ROI. Input given to the algorithm is a normalized image and output is ROI image. $P_R$, $P_G$, $P_B$ - Protocol for Red, Green and Blue.

**ROI Identification**
Step 1: Start
Step 2: Input Normalized image $NI$
Step 3: Let ROI be $\emptyset$
Step 4: For $i=0$ to Width size of $NI$ do
Step 5: For $j=0$ to Height of $NI$ do
Step 6: $PSIGN = NI \text{ (ij) RGB}$
Step 7: Let $R= PSIGN >> 16 & HD$
Step 8: Let $G= PSIGN >> 8 & HD$
Step 9: Let $B= PSIGN >> 0 & HD$
Step 10: IF ($R>PR$)
Step 11: IF ($G>PG1 \text{ AND } G<PG2$)
Step 12: IF ($B>PB1 \text{ AND } B<PB2$)
Step 13: Then ROIIJ $\rightarrow (0,0,0)$
Step 14: END IF of Step 12
Step 15: END IF of Step 11
Step 16: END IF of Step 10
Step 17: ELSE
Step 18: Then ROIIJ $\rightarrow (255,255,255)$
Step 19: END ELSE
Step 20: End for Step 5
Step 21: Return ROI image
Step 22: End

Region of interest (ROI) extraction is used to extract edges of the image along with neighbouring edges of image. Input given to algorithm is a normalized image and output is ROI image sub image.

**ROI Extraction**
Step 1: Start
Step 2: Input Normalized image $NI$, ROI Image $ROI$
Step 3: Start
Step 4: Let Flag= FALSE
Step 5: For $i=0$ to Width size of ROI do
Step 6: For $j=0$ to Height of ROI do
Step 7: $PSIGN = NI \text{ (ij) RGB}$
Step 8: Let $R= PSIGN >> 16 \text{ & HD}$
Step 9: Let $G= PSIGN >> 8 \text{ & HD}$
Step 10: Let $B= PSIGN >> 0 \text{ & HD}$
Step 11: IF value of red,green and blue pixel is not equal to 255 then
Step 12: PLU $\rightarrow I, J$ [ Left Upper Pixel]
Step 13: PRL $\rightarrow I, J$ [ Right Lower Pixel]
Step 14: Flag=TRUE
Step 15: BREAK
Step 16: End for Step 6
Step 17: IF (Flag =TRUE) then
Step 18: BREAK
Step 19: End for Step 5
Step 20: ROISUB= resize ($NI$, PLU, PRL)
Step 21: Return ROI Sub Image $ROISUB$
Step 22: End

**Step 3: Edge Traversing**
In Edge traversing, edges of the images are traversed and their luminance value is noted. If the image is real then luminance value of the edges remains intact. On the other hand if the image is fake then luminance value of the edges keep fluctuating. Edge traversing algorithm is used for this purpose.

**Algorithm :Edge Traversing**
Input: Original image and edge image.
Output: Pixel List
Step 1:Start
Step 2: For $i=0$ to width of edge image
Step 3: For $j=0$ to height of edge image
Step 4: Right Shifting signed integer by 16 bits and performing AND operation with hexadecimal number to obtain value of red.
Step 5: Right Shifting signed integer by 8 bits and performing AND operation with hexadecimal number to obtain value of green.
Step 6: Right Shifting signed integer by 0 bits and performing AND operation with hexadecimal number to obtain value of blue.
Step 7: If value of $r,g,b$ is 255 then record luminance of original image.
Step 8: Add luminance value to temporary list and later display pixel list.
Step 9: Stop

**Step 4: Edge Traversing**
While traversing the edges of images their luminance value is recorded. Formula for calculating luminance is given below. Luminance is amount of white light or brightness of the pixel
Neetu Pillai

Luminance=0.299*R + 0.587 * G + 0.114 * B

**Step 5: Luminance Correlation**
In Luminance Correlation a correlation array is constructed by fetching values of histogram and luminance. Based on the values of trained image another correlation array is created. Both the arrays are given as input to Pearson Correlation for checking the correlation between both arrays and outputs the value in the range of 0-1.

**Step 6: Fuzzy Classification**
The output of Pearson Correlation is between 0-1 and is further divided into five parts. Values between -1 to -0.6 is classified as very low. Values between -0.6 to -0.2 is classified as low. Values between -0.2 to 0.2 is classified as medium. Values between 0.2 to 0.6 is classified as high. Values between 0.6 to 1 is classified as very high. If the value is near to -1 it is fake whereas if the value is near to 1 it is real.

**4. Results and discussions**
The proposed model uses features like CNN and histogram for fake image detection along with fuzzy classifier to enhance the performance. The system is developed using Java language and NetBeans. Histogram equalization result (HTER) is used for performance analysis of fake image detection processes using CNN along with the existing methodology– FCID-HIST and FCID-FE.

FCID-HIST, FCID-FE and FCID-CNN performs cross validation for performance analysis of fake image detection processes. Below Table I consist of FCID-HIST, FCID-FE and FCID-CNN cross validation results. Average HTER of FCID-HIST is 18.42%, average HTER of FCID-FE is around 16.91%, whereas average FCID-CCN is about 14.06%. (Figure 2 and Figure 3)

<table>
<thead>
<tr>
<th>Folder no</th>
<th>FCID-CNN</th>
<th>FCID-HIST</th>
<th>FCID-FE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>14</td>
<td>21.8</td>
<td>16.65</td>
</tr>
<tr>
<td>2</td>
<td>13.9</td>
<td>22.9</td>
<td>16.9</td>
</tr>
<tr>
<td>3</td>
<td>14.1</td>
<td>21.7</td>
<td>16.65</td>
</tr>
<tr>
<td>4</td>
<td>14</td>
<td>15.2</td>
<td>16.9</td>
</tr>
<tr>
<td>5</td>
<td>14.2</td>
<td>16.8</td>
<td>16.9</td>
</tr>
<tr>
<td>6</td>
<td>13.8</td>
<td>16.65</td>
<td>17.21</td>
</tr>
<tr>
<td>7</td>
<td>14.3</td>
<td>17.78</td>
<td>16.84</td>
</tr>
<tr>
<td>8</td>
<td>14.1</td>
<td>17.59</td>
<td>17.24</td>
</tr>
<tr>
<td>9</td>
<td>14</td>
<td>16.98</td>
<td>16.89</td>
</tr>
<tr>
<td>10</td>
<td>14.2</td>
<td>16.84</td>
<td>16.98</td>
</tr>
</tbody>
</table>

![Figure 2](image.png)

*Figure 2* HTER cross validation between FCID-CNN and FCID-HIST
**Figure 3** HTER cross validation between FCID-CNN and FCID-FE

**Figure 4** Example of real and fake image

**Figure 5** Fake image detection process

*Figure 4* depicts an example of fake image and real image. *Figure 5* depicts the stepwise process of fake image detection. First the system accepts input image, later light of the image is corrected for better processing, later features like dark channel, bright channel, alpha channel, hue, saturation and histogram, Later edge of the images are extracted and luminance of the edges is recorded. If luminance of edges keeps changing it means the image is fake whereas if the luminance over the edges remains constant it means the image is real.

### 5. Conclusion

This research paper mainly focuses on fake image detection process using FCID-CNN and how it's better than the existing systems. This paper properly
Neetu Pillai

scrutinizes the different features of the images like hue, saturation, Dark channel, bright channel, an alpha channel. Then Distribution factors of all these features are being estimated using a Gaussian distribution model. Then, based on this distribution factor the convolution neural network efficiently handles the fake image detection process. FCID-CNN performs 4.36% better than FCID-HIST and 2.85% better than FCID-FE. Later on, this phony picture discovery procedure can be upgraded to all the more precisely by considering progressively features like DCT and wavelet change strategies to deal with all the more profound high definition pictures of the room and so forth.

**Acknowledgment**
None.

**Conflicts of interest**
The authors have no conflicts of interest to declare.

**References**

Neetu Pillai is Assistant Professor at Pillai HOC College of Engineering and Technology, India. She received her M.E degree in Computer Engineering from Mumbai University in 2019 and B.E degree in Computer Engineering from Mumbai University in 2017. Her key area of interest includes Image Processing, Cyber Security and Data Warehousing and Data Mining.

Email: neetu.pillai012@gmail.com