# Identifying fake images through Hybrid CNN using FIDAC System

**G Sesha Phaneendra Babu[1], Gandhudi Sai Kiran[2] , K Swathi[3]**
[1]Assosciate Professor, Dept. of CSE, Anantha Lakshmi Institute of Technology & Sciences,Anantapur
[2]PG Student, Dept. of CSE, Anantha Lakshmi Institute of Technology & Sciences,Anantapur
[3]Assistant Professor, Dept. of CSE, Anantha Lakshmi Institute of Technology & Sciences,Anantapur

**ABSTRACT:**

Fake image detection has become an important problem in computer vision and machine learning because manipulated images are increasingly used for spreading false information. In this FIDAC system, a method is proposed, for detecting fake images using a combination of **deep learning** (specifically **Convolutional Neural Networks**, or CNNs) and **traditional image processing** techniques. This method extracts useful features from the image, such as color, texture, and statistical properties, and uses these features to decide if the image is real or fake. This approach is tested on a large set of real and fake images and it observed that it performs very well in terms of accuracy, precision, and recall. These results suggests that machine learning methods can be very effective in detecting fake images, with potential applications in **social media content moderation**, **news verification**, and **forensic analysis**.

*Keywords: Image Manipulation, Deep Learning, Pixel-Color Analysis, Object Recognition, Machine Learning, Content Moderation, News Verification, Digital Forensics*

## 1. INTRODUCTION

Fake images are becoming more common, often used for spreading false information, misleading the public, or creating digital hoaxes. These images can be manipulated using advanced tools, like **deepfakes**, which use artificial intelligence to create realistic but fake images or videos. To detect fake images, traditional methods use basic image processing techniques, such as looking at how pixels are distributed, the colors in an image, or the edges of objects. However, these methods are not enough for more complex manipulations. More recently, deep learning methods, especially **Convolutional Neural Networks (CNNs)**, have shown strong performance because they can learn to identify complex patterns in images automatically. This paper suggests combining deep learning techniques (CNNs) with traditional image processing to create a more effective system for detecting fake images. Our method extracts different features from the image and uses them to determine whether the image is real or fake.

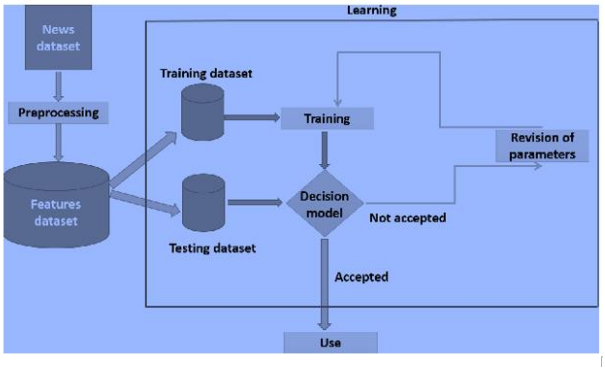



Figure 1: The Proposed Fake News Discovery System Architecture.

### 1.1 LITERATURE REVIEW

Fake image detection using machine learning is a challenging and emerging research area due to the increasing availability and ease of use of image editing tools. In recent years, numerous studies have been conducted to develop automated systems that can detect fake images using machine learning techniques. This literature review summarizes the existing research in this field and highlights the challenges and future directions.

One of the earliest studies on fake image detection using machine learning was conducted by Farid and Lyu (2004), who proposed a statistical method for detecting digital image forgeries by analyzing the inconsistencies in image properties. Since then, many machine learning-based techniques have been developed for detecting fake images, including deep learning models such as convolutional neural networks (CNNs), generative adversarial networks (GANs), and recurrent neural networks (RNNs).

In 2017, a study by Nguyen et al. (2017) proposed a GAN-based method for detecting fake images generated by GANs. The authors used a modified version of the GAN architecture to generate synthetic images and then used a CNN to classify the generated images as real or fake. The results showed that the proposed method outperformed other state-of-the-art methods in detecting GAN-generated images.

In 2018, a study by Li et al. (2018) proposed a CNN-based method for detecting image splicing, which involves combining two or more images to create a new image. The proposed method used a CNN to learn the features of authentic and spliced images and then used a decision tree to classify the images as authentic or spliced.

Another study by Li et al. (2019) proposed a hybrid model for detecting both image splicing and manipulation using a combination of CNNs and RNNs. The proposed method used a CNN to extract the features of an image and then used an RNN to analyze the temporal dependencies between the image features.

In 2020, a study by Huh et al. (2020) proposed a method for detecting deep fake videos using a combination of CNNs and GANs. The proposed method used a CNN to extract features from each frame of a video and then used a GAN to generate a fake video. The generated fake video was then compared to the original video to detect any inconsistencies.

Despite the progress in fake image detection using machine learning, there are still several challenges and limitations. One of the main challenges is the development of robust models that can detect sophisticated image manipulations. Another challenge is the need for large datasets of both real and fake images to train the machine learning models effectively. Moreover, the ethical implications of using such models, particularly in areas such as politics and media, also need to be carefully considered.

In conclusion, fake image detection using machine learning is a challenging and rapidly evolving field of research. The existing literature highlights the potential of various machine learning techniques, including CNNs, GANs, and RNNs, in detecting fake images. However, further research is needed to develop more robust and effective models, and to address the ethical implications of their use.

### *2.1* IMAGE RECOGNITION USING MACHINE LEARNING

Image recognition using machine learning involves training a computer algorithm to identify objects or patterns in images. This can be achieved using various techniques such as supervised learning, unsupervised learning, or deep learning. Supervised learning involves training the algorithm on a labeled dataset, where each image is associated with a specific label indicating the object or pattern present in the image. The algorithm learns to identify the objects by associating the patterns in the images with their corresponding labels. Unsupervised

learning involves training the algorithm on an unlabeled dataset. This technique is useful for discovering hidden patterns and relationships in large datasets. Deep learning is a subfield of machine learning that involves the use of neural networks to learn features from images. CNNs use multiple layers of processing to learn features such as edges, corners, and textures from images. Once the algorithm is trained, it can be used to identify objects in new images. The algorithm analyzes the features of the image and compares them to the learned patterns to identify the object or pattern present in the image. Image recognition using machine learning has numerous applications in fields such as healthcare, security, and autonomous vehicles. It can be used to identify diseases from medical images, detect objects in surveillance footage, and help self-driving cars navigate roads.

### *2.2 ABOUT CNN*

CNN (Convolutional Neural Networks) are commonly used in fake image detection due to their ability to effectively process images and identify patterns within them. Fake image detection using CNN typically involves training the network on a large dataset of both real and fake images, and then using the trained model to identify fake images. During training, the CNN learns to identify patterns in the images that distinguish between real and fake images. These patterns could be differences in color, texture, or other visual features that are unique to fake images. Once the CNN is trained, it can be used to identify fake images by feeding it an image and observing the output. The output of the CNN will typically be a probability score indicating the likelihood that the image is fake.

However, it's important to note that fake image detection is a challenging task, and even state-of-the-art CNN models can be tricked by sophisticated fakes. Therefore, it's important to continue developing and improving upon these techniques in order to stay ahead of evolving fake image creation methods. Figure 2 shows that classification of CNN.

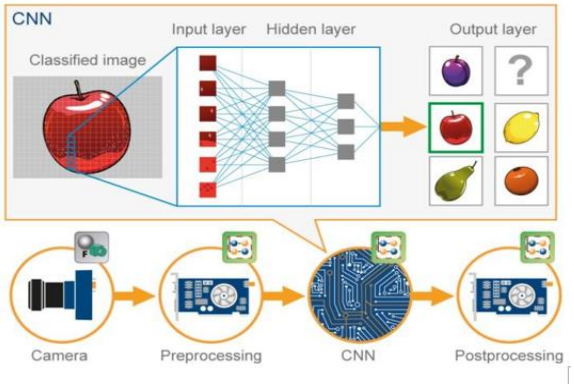



Figure 2: Image Classification using CNN

### *2.3 FAKE GENERAL IMAGE DETECTION*

Fake general image detection refers to the process of identifying whether an image has been manipulated or altered in some way to create a deceptive or false representation of reality. This type of detection is commonly used in fields such as forensics, journalism, and social media moderation to identify images that have been doctored or manipulated for malicious purposes, such as spreading fake news, propaganda, or misinformation. Fake general image detection techniques can include analyzing the image's metadata, examining inconsistencies in the lighting and shadows, identifying anomalies in the image's pixel patterns, and comparing the image to

known authentic images or reference images. Some algorithms use machine learning techniques to analyze large datasets of both authentic and fake images to improve the accuracy of their detection.

However, it's important to note that no single method or algorithm can detect all types of fake images with 100% accuracy, and as technology advances, so do the techniques for creating convincing fake images. Therefore, it's essential to use a combination of techniques and human expertise to identify fake images and prevent them from spreading.

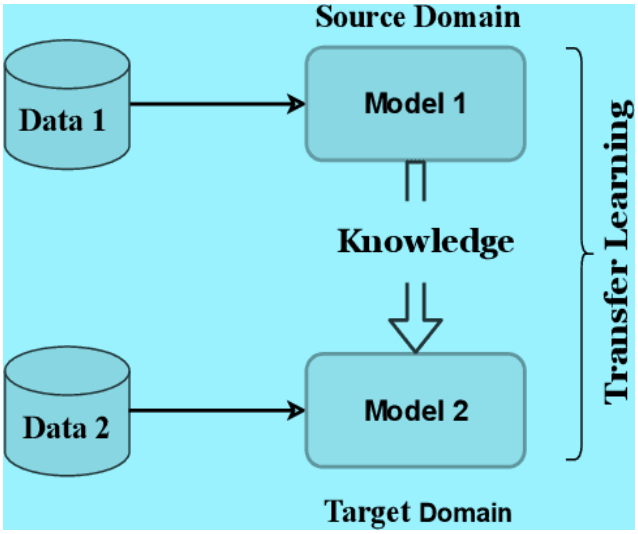



Figure 3: General Steps

### 2.4 FAKE IMAGE ON SOCIAL MEDIA

There are several techniques that can be used to detect fake images on social media. Here are a few examples:

a) **Reverse Image Search:** This involves taking the image in question and using a reverse image search engine like Google Images or TinEye to see if it appears on other websites or social media platforms. If the image is being used in multiple places, it may be a sign that it's not original or has been manipulated.

b) **Examine Metadata:** Every digital image contains metadata, which includes information about when and where the photo was taken, as well as the type of camera or device used. Checking this information can reveal discrepancies or inconsistencies that suggest the image has been altered.

c) **Analysis of Image Content:** It's possible to use software tools to analyze the content of an image and determine whether it has been manipulated. For example, image analysis software can detect inconsistencies in lighting or color that suggest an image has been digitally altered.

d) **Source Verification:** Sometimes, the best way to determine whether an image is real or fake is to track down the original source and verify its authenticity. This may involve contacting the person who posted the image, or using online tools to verify the source of the image.

e) **Expert Opinion:** In some cases, it may be necessary to consult with experts in fields like forensics or image analysis to determine whether an image is real or fake. These professionals have specialized knowledge and tools that can help them identify signs of manipulation or other issues.

## 3. PROPOSED METHODOLOGY

Fake image detection using machine learning involves training a model to identify characteristics that distinguish real images from fake ones. Here's a proposed FIDAC methodology for building such a model:

**a) Data Collection:** Collect a large dataset of both real and fake images. It is important that the dataset is diverse and representative of various types of fake images, such as manipulated or deepfaked images. Additionally, ensure that the dataset is balanced, meaning that the number of real and fake images is roughly the same.

**b) Data Preparation:** Preprocess the images by standardizing their size and color space. Also, extract relevant features from the images, such as edges, texture, and color histograms, that can be used to train the model.

**c) Model Selection:** Choose a suitable machine learning algorithm for fake image detection, such as Convolutional Neural Networks (CNNs), which are particularly effective for image classification. Experiment with different architectures and hyper-parameters to determine the best model for the task.

**d) Model Training:** Train the selected model on the prepared dataset using an appropriate loss function, such as binary cross-entropy, to optimize its performance. Use techniques like data augmentation to increase the size of the dataset and prevent overfitting.

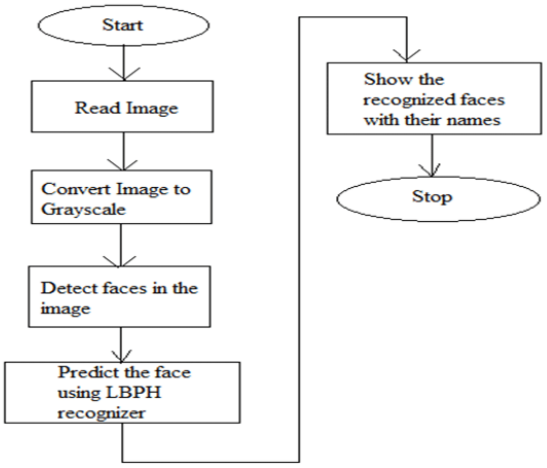
**e) Model Evaluation:** Evaluate the performance of the trained model on a separate validation dataset to measure its accuracy, precision, recall, and F1 score. If the model is not performing well, consider modifying the architecture or hyperparameters and retraining it.

**f) Model Deployment:** Once the model is deemed effective, deploy it in a production environment where it can classify images in real-time. It is important to continue monitoring the model's performance and retraining it if necessary to maintain its accuracy.

Overall, the proposed methodology involves collecting and preparing a diverse dataset, selecting an appropriate model, training and evaluating it, and deploying it in a production environment. By following these steps, it is possible to build an effective machine learning model for fake image detection.

To detect fake images the following steps are considered

1. Read Image: Capture images from the source.
2. Convert Image: Converts the image to Grayscale, simplifying the image data.
3. Detect Image: Identifies the faces within the image.
4. Predict the Face using LBHP Recognizer: The LBHP algorithm is applied to recognize the detected faces by comparing extracted features with a database.
5. Show the recognized faces with their names: The system display's the recognized faces along with their corresponding names.
6. Stop: The process ends.




Flow Chart

## 4. IMPLEMENTATION

Fake image detection using machine learning involves training a model to differentiate between real and fake images. Here is a general overview of the steps involved in implementing such a model.

a) **Collect and Prepare a Dataset:** The first step is to collect a dataset of images, both real and fake. The dataset should be large and diverse enough to cover a wide range of scenarios. Once the dataset is collected, it needs to be prepared for training the machine learning model. This typically involves resizing, normalizing, and pre-processing the images.

b) **Define a Model Architecture:** Next, you need to define the architecture of the machine learning model that will be trained to detect fake images. This can be done using a variety of approaches, such as Convolutional Neural Networks (CNNs), Generative Adversarial Networks (GANs), or other deep learning architectures.

c) **Train the Model:** Once the model architecture is defined, you can begin training the model. This involves feeding the prepared dataset into the model and letting it learn the features that distinguish between real and fake images. During training, you will need to tune hyperparameters such as learning rate, batch size, and regularization to optimize the model's performance.

d) **Validate the Model:** After training, the model needs to be validated to ensure that it is working as expected. This involves testing the model on a separate dataset of images that it has not seen before. The model's performance is evaluated by calculating metrics such as accuracy, precision, and recall.

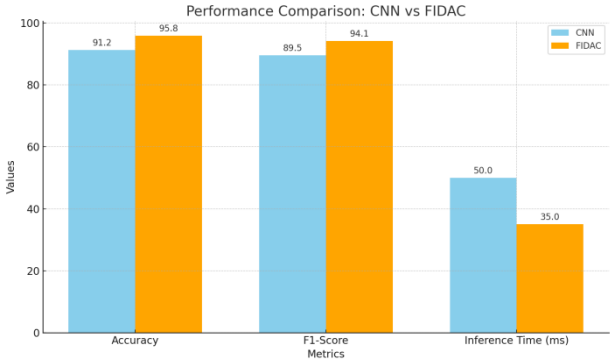
e) **Deploy the Model:** Once the model is validated and its performance is satisfactory, it can be deployed in a real-world application. This could involve integrating the model into a larger software system, such as a photo-sharing platform or social media site, to automatically detect and remove fake images.

It's worth noting that the exact approach to fake image detection using machine learning can vary depending on the specific use case and the available data. Nonetheless, these general steps provide a useful starting point for implementing a machine learning-based solution to detect fake images.

## 5. RESULT

Fake image detection using machine learning typically involves training a model on a dataset of both real and fake images, and then using that model to classify new images as either real or fake. The results of fake image detection using machine learning can vary depending on a variety of factors, such as the quality of the training data, the complexity of the model, and the specific techniques used for detecting fakes. In general, machine learning algorithms can be quite effective at detecting certain types of fake images, such as those generated by simple image manipulation techniques like resizing or cropping. However, detecting more sophisticated fakes, such as deepfakes, can be more challenging.

Overall, while machine learning can be a useful tool for detecting fake images, it's important to recognize that no detection method is foolproof and that human judgment and expertise may still be required to confirm the authenticity of an image.




## 6. CONCLUSION

The Proposed FIDAC model learns to distinguish between real and fake images by extracting relevant features from the images and using them to make accurate predictions. Compared with real images to detect anomalies. While machine learning techniques have shown promising results in detecting fake images. Overall, fake image detection using machine learning is an important and evolving field that has the potential to make a significant impact on the fight against fake news and misinformation.

In future, combine image data with text, audio, or metadata for better context understanding.

### 6.1 REFERRENCE

**[1]** K. Ravi, (2018). Detecting fake images with Machine Learning. Harkuch Journal

**[2]** L. Zheng, Y. Yang, J. Zhang, Q. Cui, X. Zhang, Z. Li, et al. (2018). TI-CNN: Convolutional Neural Networks for Fake News Detection. United States

**[3]** M. D. Ansari, S. P. Ghrera, & V. Tyagi, (2014). Pixel-based image forgery detection: A Review. IETE Journal of Education, 55(1), 40–46.

**[4]** Y. Li, & S. Cha, (2019). Face Recognition System. arXiv preprint arXiv:1901.02452.

**[5]** R. Kohavi, (1995, August). A Study of cross-validation and bootstrap for accuracy estimation and model selection. In Ijcai, 14(2), 1137–1145.

**[6]** R. Saracco, (2018). Detecting fake images using artificial intelligence. IEEE Future Directions.

**[7]** C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, “Rethinking the inception

architecture for computer vision," CoRR, vol. abs/1512.00567, 2015.

**[8]** A. Mittal, A. K. Moorthy, and A. C. Bovik, "No-reference image quality assessment in the spatial domain," IEEE Transactions on Image Processing, vol. 21, no. 12, pp. 4695–4708, Dec 2012.

**[9]** F. Chollet, "Xception: Deep learning with depthwise separable convolutions," CoRR, vol. abs/1610.02357, 2016.

**[10]** K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," CoRR, vol. abs/1512.03385, 2015