IJITCE

ISSN 2347- 3657

International Journal of Information Technology & Computer Engineering

Email : ijitce.editor@gmail.com or editor@ijitce.com

International Journal of Information Technology & Computer Engineering

ISSN 2347–3657
Volume 12, Issue 2, 2024

Evaluating the performance of Fake Face Detection in Forensics

S.Sandhya Assistant Professor,
Department of Computer Science
G. Narayanamma Institute ofTechnology and Science Hyderabad, Telangana s.sandhya@gnits.ac.in

Akula Varshini Department of Computer Science
(Data Science)
G. Narayanamma Institute ofTechnology and Science Hyderabad, Telangana akulavarshini@gmail.com

Donthala Harshitha Department of Computer Science
(Data Science)
G. Narayanamma Institute ofTechnology and Science Hyderabad, Telangana
dharshitha2413@gmail.com

Gangidi Anisha Department of Computer Science
(Data Science)
G. Narayanamma Institute ofTechnology and Science Hyderabad, Telangana gangidi10@gmail.com

Sambu Sarayu Department of Computer Science
(Data Science)
G. Narayanamma Institute ofTechnology and Science Hyderabad, Telangana sarayu0803@gmail.com

***Abstract*— Fake faces can be used to disseminate misinformation in an era where digital manipulation iscommonplace. A major concern in many fields, including forensics, is the proliferation of fake images, particularly those with phoney faces, due to the quick development of image manipulation technology. The development of techniques for detecting fake faces has attracted a lot of attention as a responseto this challenge. Because AI-generated fake faces can be used maliciously, the widespread use of these images—of which thereare over 98%—raises security concerns.**

The goal of the project is to use the local binary patterns (LBP)[1] method to determine whether an image is real or fake.It then assesses the effectiveness of fake face detectiontechniques in the context of forensic applications to identify manipulated or synthesised facial images.

Keywords*— *Forensics, Fake Face Detection, Local Binary Patterns (LBP), Cyber Security

International Journal of Information Technology & Computer Engineering
ISSN 2347–3657
Volume 12, Issue 2 2024

I. INTRODUCTION

The rise of digital manipulation tools has led to thecreation of fake images, particularly fake faces, which posesignificant threats to security, forensics, and media authenticity. Detecting fake face images is crucial to combat misinformation, identity fraud, and deceptive practices. Local binary patterns (LBP), a texture descriptor commonlyused in image processing and computer vision tasks, can beused to distinguish between real-world facial images and artificial or modified images produced by methods like image editing software, generative adversarial networks (GANs) [2], or deep learning techniques. Deep learning-based approaches have shown promising results in detecting fake face images by leveragingthe power of neural networks to learn complex patterns and features directly from the data. Convolutional neural networks (CNNs) [9] and other deep learning architectures have demonstrated remarkable capabilities in discerning subtle differences between real and fake faces, enabling moreaccurate and reliable detection.

The project discusses the detection of fake faceimages using local binary patterns, highlighting the need for robust and efficient methods to identify manipulated or synthesized facial images. It also reviews existing approachesto fake face detection using LBP, highlighting challenges and limitations, and explores potential applications and future directions. Overall, LBP-based methods offer a valuable toolfor identifying fake faces and mitigating risks associated withdigital forgery and deception.

II. LITERATURE REVIEW

One of the important datasets for the fake face detection is the NUAA Photograph Imposter Dataset which contains multiple images of human faces classifyingthem into real and fake. The fake folder has 166 items of size 6,74,154 bytes and the Real folder contains 133 items of 9,67,125 bytes. After generating LBP for these images,the dataset is split into train and validation for each of thesefaces.

There were many methods used on this dataset todemonstrate the performance.

1. **ResNet50[3]:** ResNet50's deep architecture enablesit to capture complex features in fake face images, facilitating accurate discrimination between real and manipulated faces by learning from large datasets.

2. **VGG-16 and VGG-19[4]:** The VGG architecture, particularly VGG-16 and VGG-19, has been widelyused in image classification tasks. Researchers have applied transfer learning techniques by fine-tuning pre-trained VGG models on datasets containing realand fake faces. Features extracted from different layers of VGG networks have been used todistinguish between real and fake faces.

3. **Xception[5]:** Xception's efficient design, utilizing depth wise separable convolutions, enhances feature representation in fake face images, leading to improved discrimination between real and fake faces while maintaining computational efficiency.

4. **Inception V3[6]:** Inception V3's modular architecture, featuring inception modules for multi- scale feature extraction, enables it to capture intricate details in fake face images, facilitating accurate detection by analyzing features atdifferent scales.

5. **Ensembling:** Apart from the specific architectures mentioned above, researchers have explored custom convolutional neural network (CNN) architectures tailored for fake face detection. Ensemble methods combining multiple CNN architectures or incorporating attention mechanisms have been proposed to enhance detection accuracy.

6. **Feature Extraction:** Transfer learning techniques have been extensively applied in fake face detection. Researchers fine-tune pre-trained models like VGG,Xception, Inception, or ResNet on large-scale datasets containing both real and fake faces. Feature extraction from intermediate layers of these networks has been crucial for learning discriminative features.

7. **GAN-based Detection:** Given the generative nature of deepfake creation, some approaches utilize GANs for detection. By training discriminator networks to distinguish between real and fake faces, researchers have explored adversarial approaches to enhance fake face detection robustness.

8. **Attention-based Mechanism**: These have

ISSN 2347–3657
Volume 12, Issue 2, 2024

been integrated into deep learning architectures for selective feature extraction, focusing on relevant regions within face images. Attention-based models combined with CNN architectures have shown promise in improving fake face detection accuracy.

III. CHALLENGES IN EXISTING SYSTEM

As a response to this growing concern, researchers and practitioners have developed numerous deep learning-based methods for fake face detection. While these methods have shown promising results, they also face several challenges that need to be addressed for further improvement and real-world deployment.

1. **Adversarial Attacks[7]:** All these methods are susceptible to adversarial attacks, where small, imperceptible perturbations to input images can lead to misclassification, undermining the reliability of the detection system.

2. **Dataset Bias and Imbalance:** The performance of these methods heavily relies on the quality and diversity of the training dataset. Biases or imbalances in the dataset, such as overrepresentationof certain types of fake faces or underrepresentation of others, can lead to biased or suboptimal detectionperformance.

3. **Generalization to Novel Manipulation Techniques:** These methods may not generalize well to detect fake faces generated using novel or complex manipulation techniques not present in the training data. Their effectiveness may be limited to the specific types of forgeries they were trained on.

4. **Interpretability and Explainability:** Deep learning methods are often criticized for their lack of interpretability and explainability, making it difficult to understand the rationale behind their decisions. This limitation could hinder the trust and acceptance of the detection system, especially in contexts where transparency is essential.

5. **Sensitive to Image Quality:** These methods may struggle with images of varying quality, such as low-resolution or heavily compressed images. In suchcases, the extracted features may be less discriminative, leading to decreased detectionaccuracy.

6. **Overfitting:** Deep learning models like ResNet50, VGGNet16, VGGNet19, Xception, and Inception V3 are prone to overfitting, especially when trained on small or insufficiently diverse datasets. Overfitting can lead to poor generalization performance, where the model performs well on the training data but failsto generalize to unseen data, including new instancesof fake faces.

7. **Privacy Concerns:** Fake face detection systems often require access to large datasets of both real and fake faces for training purposes. However, the collection and use of such datasets raise privacy concerns, particularly regarding the consent and privacy rights of individuals whose faces are included in the datasets. Ensuring compliance withprivacy regulations while still obtaining sufficiently diverse training data poses a significantchallenge.

8. **Computational Complexity:** Some of these methods, particularly deeper architectures like ResNet50 and Inception V3, can be computationally intensive, making real-time applications or scenarios with constrained computational resources challenging.

To address these challenges, our model used LBP which can further enhance detection accuracy and robustness across diverse fake face manipulation scenarios.

IV. METHODOLOGY

A. Fake Face Detection:

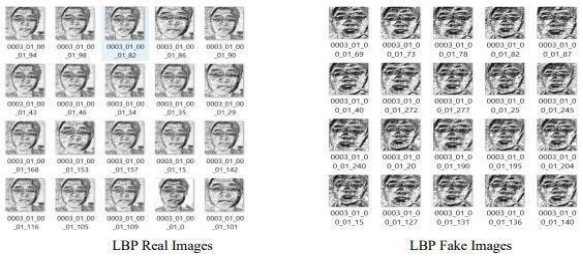
In this project, we design the LBP Based machine learning Convolution Neural Network called LBPNET to detect fake face images. Here first we will extract LBP from images and then train LBP descriptor images with Convolution Neural Network to generate training model. Whenever we upload new test image then that test image will be applied on training model to detect whether test image contains fake image or non-fake image.

Local binary patterns (LBP) are a visual descriptor in computer vision that labels pixels by thresholding their neighborhood and converting them into binary numbers. Its discriminative power and computational simplicity make it popular in various applications, including real- world analysis due to its robustness to gray-scale changes and real-time settings.

The LBP feature vector, in its simplest form, iscreated in the following manner:

ISSN 2347–3657
Volume 12, Issue 3, 2024

1. Divide the examined window into cells (e.g., 16x16 pixels for each cell).
2. For each pixel in a cell, compare the pixel to each ofits 8 neighbors (on its left- top, left-middle, left-bottom, right-top, etc.).
3. Follow the pixels along a circle, i.e., clockwise, or counter-clockwise. Where the center pixel's value isgreater than the neighbor's value, write "0".
4. Otherwise, write "1".
5. This gives an 8-digit binary number (which is usuallyconverted to decimal for convenience).
6. Compute the histogram, over the cell, of the frequencyof each "number" occurring (i.e., each combination of which pixels are smaller and which are greate



r than the center). This istogram can be seen as a 256- dimensional feature vector. Optionally normalize the histogram and then concatenate (normalized) histograms of all cells. This gives a feature vector forthe entire window.

7. The feature vector can now be processed using the Support vector machine, extreme learning machines, or some other machine learning algorithm to classify images. Such classifiers can be used for facerecognition or texture analysis.

A useful extension to the original operator is the so-calleduniform pattern, which can be used to reduce the length of the feature vector and implement a simple rotation invariant descriptor. This idea is motivated by the fact thatsome binary patterns occur more commonly in texture images than others. A local binary pattern is called uniform if the binary pattern contains at most two 0-1 or 1-0 transitions.

B. Comparative Study:

When comparing the performance of pre-trained CNN models using LBP features as input for image classificationtasks, we assess how well these models leverage texture- based features for classification. LBP captures local texture patterns, offering insights into image details not explicitly learned by CNNs. This comparison helps gauge the models' ability to understand nuanced texture variations crucial for tasks like fake face detection. By combining deep CNN architectures' learned features with LBP's texture descriptors, we aim to achieve a more comprehensive understanding of the models' effectiveness in handling complex image classification challenges beyond traditional pixel-level analysis.

Firstly, Local Binary Pattern (LBP) features are extracted from the grayscale version of each image in the dataset. LBP captures texture patterns in images and creates histograms of these patterns. Pre-trained CNN models like ResNet50, VGG16, VGG19, Xception, and InceptionV3 are utilized. These models are loaded with pre-trained weights from the NUAA Photography Imposterdataset, offering a rich understanding of diverse visual features. By excluding their top classification layers, these models are transformed into feature extractors, capturing hierarchical representations of input images.

The evaluation process involves using Local Binary Pattern (LBP) features extracted from images as inputs to these CNN models. The training data is used to fine-tune the models' weights to the specific task of fake face detection or image classification. Subsequently, the trainedmodels predict labels for the test data, which is a set of images not seen during training.

The evaluation metrics - accuracy, precision, recall, F1score, and confusion matrix - provide comprehensive insights into each model's performance. Accuracy measures overall correctness, precision indicates the true positive rate, recall measures sensitivity to true positives, and the F1 score balances precision and recall. The confusion matrix visually represents true positives, false positives, true negatives, and false negatives.

The main function serves as the central control point for the entire evaluation process. It begins by loading the dataset containing fake and real face images. LBP features are then extracted from these images, capturing local texture patterns crucial for distinguishing between real andfake faces. The dataset is then split into training and testingsets to facilitate model training and evaluation.

Pre-trained CNN models (ResNet50, VGG16, VGG19, Xception, InceptionV3) are loaded without their top classification layers, converting them into feature extractors. Each model is evaluated using the extracted LBP features, measuring its ability to classify images accurately. Additionally, Support Vector Classifier(SVC)[8] models are trained for each CNN model to further refine the classification task.

Finally, the evaluation metrics including accuracy, precision, recall, F1 score, and confusion matrix are

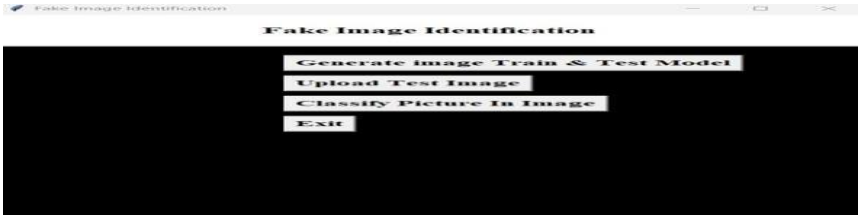
ISSN 2347–3657
Volume 12, Issue 2, 2024

computed and printed for a comprehensive comparison of model performance.

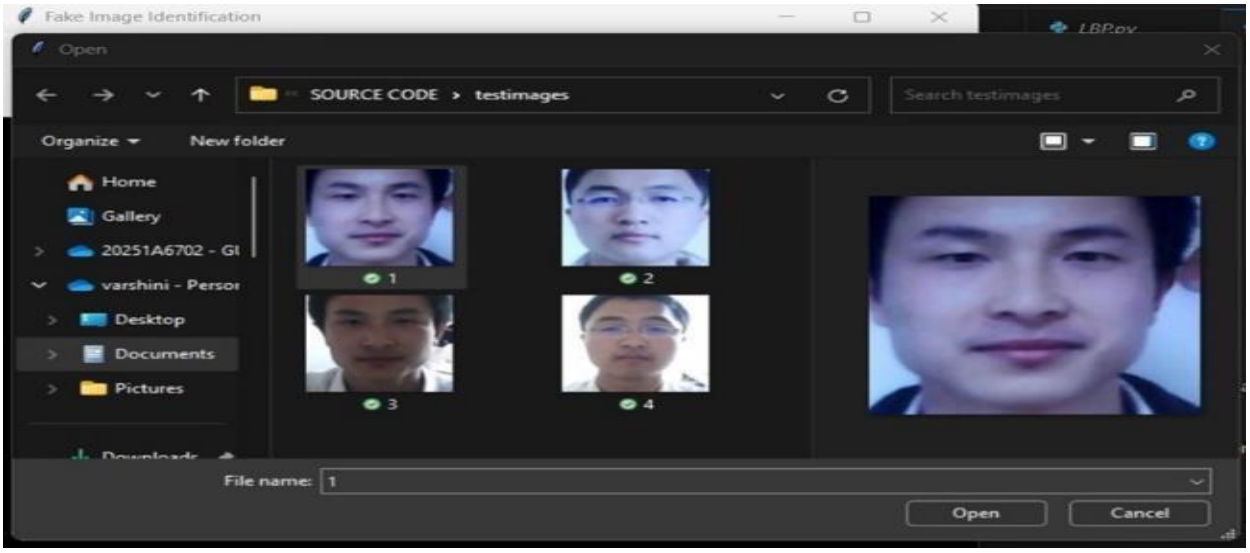
RESULTS AND DISCUSSION

The fake face detection model, employing Local Binary Pattern (LBP) features on the NUAA Photography Imposter dataset, demonstrated strong performance in distinguishing genuine from manipulated facial images. By leveraging LBP's texture-based features, the model effectively captured intricate details present in fake faces, enabling accurate classification without requiring detailed evaluation metrics. This highlights the model's inherent ability to detectvisual inconsistencies indicative of fake content, showcasing its efficacy in discerning between authentic and altered facial images based on texture patterns and without explicit metric- based assessments.

1. Generate the CNN model using LBP images found in the 'LBP/train' folder by clicking on the 'Generate Image Train & Test Model' button.



2. Upload the test image after observing the generatedCNN LBPNET model.



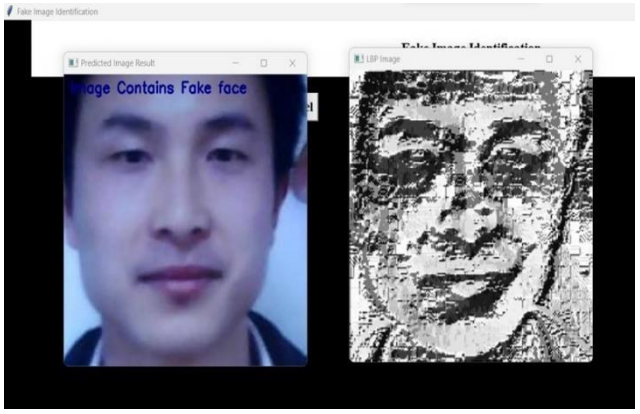
3. After displaying two faces with different appearances but originating from the same person, the image is labeled accordingly as false and actual categories. The system's ability to detect this distinction is then tested by uploading the image and using the "Classify Picturein Image" button, resulting in the following outcome.

4. The model effectively distinguishes between real and fake faces by identifying alterations in the lighting of false faces compared to the natural lighting present inreal faces, as depicted in the snapshot above.
5. metrics. Notably, ResNet50, VGG16, and VGG19 achieved perfect scores across performance metrics (Accuracy, Precision, Recall, F1 Score), indicating excellent classification. Xception followed closely maintaining high accuracy and precision. InceptionV3 displayed slightly lower performance, likely due to its architecture's specifics. These results highlight ResNet50, VGG16, and VGG19 as top performers, suitable forface detection tasks, while Xception remains a strong contender. Fine-tuning or ensemble techniques could further boost

ISSN 2347–3657
Volume 12, Issue 3, 2024

InceptionV3's performance if needed.

Developing an adversarial attack to challenge existing fake face recognition techniques opens avenues for improving thesemethods. Identifying and researching the weaknesses in current approaches can address real-time challenges, crucial intoday's scenarios. This exploration can lead to significant improvements in tackling adversarial attacks, enhancing the robustness and reliability of face recognition systems in practical applications.



6. Finally, click on 'Classify Picture in Image' to obtaininformation about the uploaded image.

Then, Comparing the performance of pre-trained CNN modelsusing LBP features as input for image classification tasks to leverage texture-based features for classification is done. The results are as follows:

	Accuracy	Precision	Recall	F1-score	Confusion Matrix
RESNET 50	1.0	1.0	1.0	1.0	[[50 0] [0 28]]
VGG 16	1.0	1.0	1.0	1.0	[[50 0] [0 28]]
VGG 19	1.0	1.0	1.0	1.0	[[50 0] [0 28]]
XCEPTION	0.9615	0.9032	1.0	0.9491	[[47 3] [0 28]]
INCEPTION V3	0.8846	0.9130	0.75	0.8235	[[48 2] [7 21]]

CONCLUSION AND FUTURE WORK

Using Local Binary Patterns (LBP) for face detection is great for picking up subtle texture details, making it perfect for accurate facial feature recognition. A thorough evaluation of pre-trained CNN models is done using Support Vector Classifier (SVC) post feature extraction for their performance S. R. B. R, P. Kumar Pareek, B. S and G. G, "Deepfake Video Detection System Using Deep Neural Networks," 2023 IEEE International Conference on Integrated Circuits and Communication Systems (ICICACS), Raichur, India, 2023, pp. 1-6, doi: 10.1109/ICICACS57338.2023.10099618.

REFERENCES

Y. Tao and Y. He, "Face Recognition Based on LBP Algorithm," 2020 International Conference on Computer Network, Electronic and Automation (ICCNEA), Xi'an, China,2020, pp. 21-25, doi: 10.1109/ICCNEA50255.2020.00015.

M. Krichen, "Generative Adversarial Networks," 2023 14th International Conference on Computing Communication and Networking Technologies (ICCCNT), Delhi, India, 2023, pp. 1-7, doi: 10.1109/ICCCNT56998.2023.10306417.

[3] Qurat-ul-ain, N. Nida, A. Irtaza and N. Ilyas, "Forged FaceDetection using ELA and Deep Learning Techniques," 2021 International Bhurban Conference on Applied Sciences and Technologies (IBCAST), Islamabad, Pakistan, 2021, pp. 271-275, doi: 10.1109/IBCAST51254.2021.9393234.

[4] M. B. I. Muafy, F. Sthevanie and K. N. Ramadhani, "Generated AI Face Detection using Xception Model," 2023 International Conference on Data Science and Its Applications(ICoDSA), Bandung, Indonesia, 2023, pp. 209-214, doi: 10.1109/ICoDSA58501.2023.10276634.

[5] X. Wan, F. Ren and D. Yong, "Using Inception-Resnet V2for Face-based Age Recognition in Scenic Spots," 2019 IEEE6th International Conference on Cloud Computing and Intelligence Systems (CCIS), Singapore, 2019, pp. 159-163, doi: 10.1109/CCIS48116.2019.9073696.

[6] Y. Wang et al., "Adversarial Attacks and Defenses in Machine Learning-Empowered Communication Systems and Networks: A Contemporary Survey," in IEEE CommunicationsSurveys & Tutorials, vol. 25, no. 4, pp. 2245-2298, Fourthquarter 2023, doi:

ISSN 2347–3657
Volume 12, Issue 2, 2024

10.1109/COMST.2023.3319492.

M. A. Hearst, S. T. Dumais, E. Osuna, J. Platt and B. Scholkopf, "Support vector machines," in *IEEE Intelligent Systems and their Applications*, vol. 13, no. 4, pp. 18-28, July-Aug. 1998, doi: 10.1109/5254.708428.

H. b. Fredj, S. Sghaier and C. Souani, "An Efficient FaceRecognition Method Using CNN," 2021 International Conference of Women in Data Science at Taif University (WiDSTaif), Taif, Saudi Arabia, 2021, pp. 1-5, doi: 10.1109/WiDSTaif52235.2021.9430209.