JCBI
Journal of Computing & Biomedical Informatics
ISSN: 2710 - 1606

Volume 05 Issue 01
2023

Research Article
https://doi.org/10.56979/501/2023

Implementation of Efficient Deep Fake Detection Technique on Videos Dataset Using Deep Learning Method

Muhammad Mussadiq Rafiqee[1], Zahid Hussain Qaiser[1], Muhammad Fuzail[1], Naeem Aslam[1], and Muhammad Sajid Maqbool[2*]

[1]Department of Computer Science, NFC Institute of Engineering and Technology Multan, Pakistan.
[2]Department of Computer Science, Bahauddin Zakariya University,Multan, Pakistan.
[*]Corresponding Author: Muhammad Sajid Maqbool. Email: sajidmaqbool7638@gmail.com.

Received: April 01, 2023 **Accepted:** June 02, 2023 **Published:** June 05, 2023.

Abstract: Deep fake technology has recently made tremendous advances that have made it possible to produce incredibly realistic fake audio, video, and image media. These materials present serious difficulties for people. Impersonation, false information, or even a national security danger, could compromise authentication. There is currently an arms race between deep fake creators and deep fake detectors as a result of the several deep fake detection algorithms that have been suggested to keep up with these rapid advancements. But these detectors are typically unreliable and frequently miss deep fakes. This study emphasizes to suggest a machine learning technique to minimize the difficulties they encounter in identifying deep fakes in videos dataset. DL and ML is used in this study to proposed a neural network for detection deep fakes from the videos. This study initially selects a video dataset from the from the well-known Kaggle dataset repository. Secondly this dataset is augmented into two classes, real videos and fake videos, and the dataset is divided into training and testing. Thirdly, the preprocessing of dataset is done by face-extraction, region of interest selection and frames extraction to detect the real and fake videos. Fourthly, Neural Network is applied on the processed dataset and evaluate the model by calculated the accuracy. Finally, Proposed model is compare with the other models such as (Resents, Inception V3 and vision transformers). The comparison shows that our proposed model is perform well on the processed dataset as compared to other models and achieved accuracy of 94.86 percent.

Keywords: Machine Learning, Deep Learning, Deep_Fake Detection, Videos Prediction, Video Framing.

1. Introduction

Since film and digital visuals have a strong impact on both people and societal discourse, switching faces in pictures has a long history that dates back more than 150 years. Until recently, altering videos or creating phony yet convincing images required specialized knowledge or expensive computing resources. Deep fakes, a new technology that can create incredibly realistic face swapped videos, has just come to light [1]. Beyond a consumer-grade GPU, no specialized gear is needed to produce a deep fake, and various commercial software programs for the task have been made available. Their use in creating parody videos for entertainment and for use in targeted attacks against people or institutions has increased dramatically as a result of the interaction of these elements. With the knowledge that anyone can now easily produce convincing phony face-swapped videos with little gear, the requirement for developing automated detection systems becomes apparent. Digital forensics professionals can examine a single, powerful video for signs of manipulation, but they are unable to assess each of the hundreds of thousands of videos that are submitted to social media or the Internet every day. Scalable techniques are required to detect deep fakes at a large scale, and computer vision or multimodal models are particularly well adapted for this task. Although it is simple to produce multiple plausible Deep fakes, the expense of creating the hundreds of thousands of deep fake movies required to train these models is frequently too high. As a result, these models need training data. We have created and made available the largest deep fake detection dataset to date in order to hasten improvements in the state of the art of Deep fake detection [2].

Deep fakes are made utilizing DL technology to change images and videos in a way that makes it difficult for viewers to tell the difference between authentic and fake content. Deep learning has proven to be quite successful at identifying DF. Numerous studies have been conducted in recent years to learn more about the inner workings of deep fakes, and academics have created a number of deep learning-based methods to recognize deep fake films and images [3]. The goal of AI, a subfield of computer science, is to build intelligent machines that can mimic human actions and reactions. The creation of technology that enables intelligent machine and computer operations is frequently the primary goal of AI research. Usually, it is specialized and technical. Planning, manipulation, and motion, problem-solving, general and social intelligence, logical thinking, information representation, and NLP have all been the subject of recent research [4]. Robotics, scheduling, data mining, logistics, video games, healthcare, automotive, government, finance, and economics are among the sectors that benefit from AI techniques. ANN that can learn and reason using algorithms are used in DL. The term "deep" refers to the amount of layers included in the network, and these neural networks are composed of numerous layers of nodes between the input and output layers. Deeper networks allow for the identification of more complex features by utilizing several intermediate layers. DL has become essential for analyzing large volumes of data, handling complex algorithms, achieving high performance with big data sets, and extracting meaningful features [5]. The overall flow for detecting deep fakes in images and videos can be seen in Figure 1 (Figure 1 provides an illustration of the detection process for deep fakes in both images and videos).

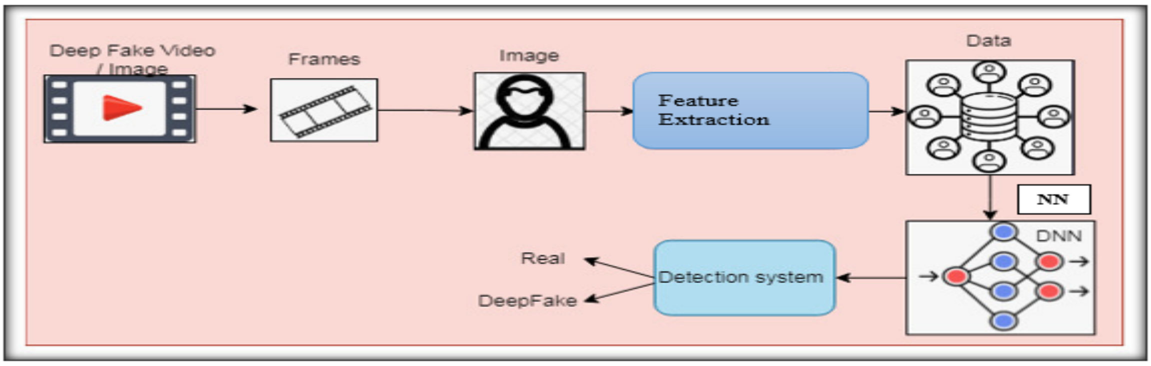

Figure 1. General flow of deep fakes detection system

2. Related Work

We discuss some relevant studies in deep fakes, machine learning, and deep learning in this chapter. This chapter describes a variety of strategies that academics have used to build models to find deep fakes and talks about the development of ML, DL, and DF.

In this age of fake news, unfettered access to vast public databases combined with the rapid advancement of DL techniques, particularly Generative Adversarial Networks, has resulted in the creation of extremely realistic fake material with related ramifications for society [1,3,6]. Recent developments in machine learning and social media platforms make it easier to produce and quickly distribute convincing fake content (such as photos, videos, and audios). Initially, fake content was created by manipulating either audio or video streams, but these days, both audio and visual streams are altered to create more convincing deepfakes. Deepfakes detection researchers primarily concentrate on identifying fake videos that only use audio or visual modalities [2].

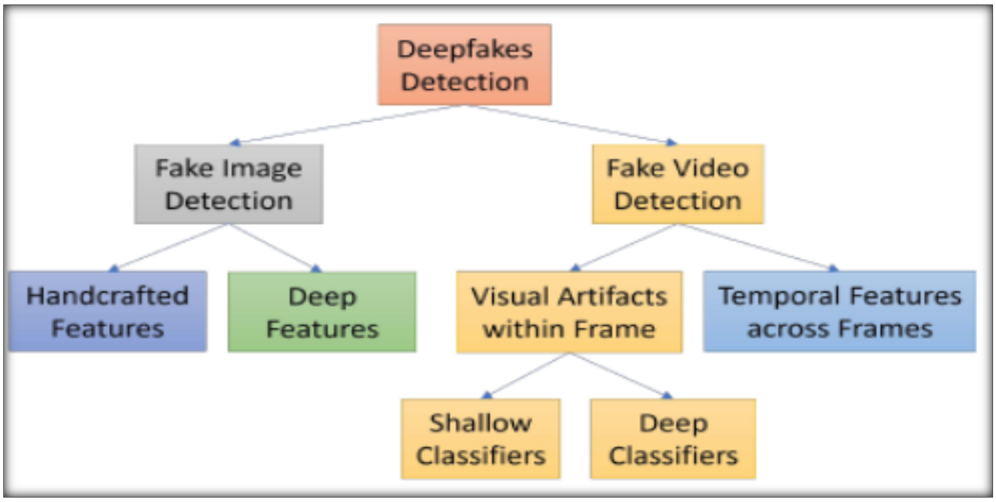

Figure 2. Categories of Deep fake Techniques

Ruben Tolosana et al [1] conduct a review paper on the deep fake domain and This survey offers a comprehensive overview of face picture modification technologies, including deep fake approaches, as well as tools for spotting such manipulations. Reviewing four distinct facial manipulation techniques in particular: complete face synthesis, identity swap (DeepFakes), attribute manipulation, expression swap, and whole-face synthesis. We give specifics on manipulation methods, currently accessible public databases, important benchmarks for technical evaluations of false detection systems, and a summary of the findings from those evaluations for each manipulation group. We give particular attention to the most recent DeepFakes generation among all the topics covered in the survey, stressing its advancements and difficulties with false identification.

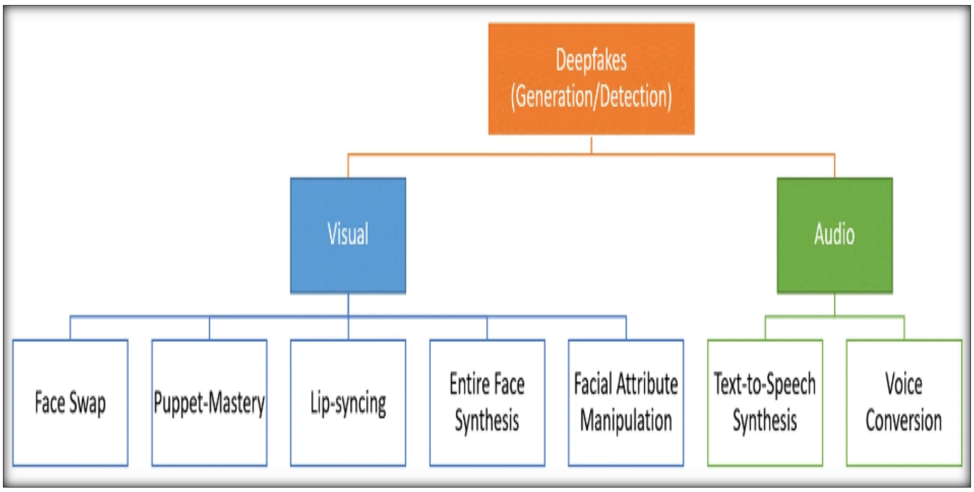

Figure 3. Deep Fake detection and generation

Hafsa Ilyas et al. [2] conducted a study on detecting deep fakes in audios, highlighting the limitations of existing methods that are rarely tested on multimodal datasets where modifications occur in both audio and visual streams. They introduced a novel architecture called AVFakeNet, which focuses on both the audio and visual modalities of a video. AVFakeNet is a unified framework based on DST-Net, comprising input, feature extraction, and output blocks. The feature extraction block utilizes a customized swim transformer module, while the input and output blocks consist of dense layers. The researchers evaluated the efficacy of their framework on five distinct datasets containing audio, visual, and audio-visual deep fakes, such as FakeAVCeleb,Celeb-DF,ASVSpoof-2019LA, World Leaders dataset, and Presidential Deepfakes dataset". Experimental results demonstrated the framework's ability to accurately identify deep fake videos by examining both the audio and visual streams. The contributions of their work include the introduction of AVFakeNet as a unified framework, enhancing the performance of deep fake detection with Dense Swin Transformer Net, resilience to high-quality deep fakes with various poses, lighting, racial makeup, gender, and age groupings, extensive testing on diverse datasets, and improved generalization compared to existing models.

Pummy Dhiman et al. [3] presented a DL technique to detect deep fakes in news articles. The study aimed to identify disinformation using deep learning and conducted a scientometric analysis of 569 papers published in the Scopus database between 2012 and mid-2022. The analysis concentrated on productivity patterns, authorship and collaboration patterns, bibliographical coupling, publishing and reference structures, authorship trends, and research trends. Findings showed a rising trend in publications during 2016, highlighting the ongoing problem of incorrect information around the world. Thematic analysis highlighted specialised domains including "DF identification," "digital contents," "electronic forensics," and "computer vision," whereas studies on "fake news," "social media monitoring," and "surveillance of public views and opinions" were deemed necessary but underdeveloped. The analysis also identified China and the USA as having the greatest international cooperation, despite India contributing more articles. Furthermore, the study evaluated the state of the art in DL algorithms for fake news identification, providing researchers with a potential roadmap. The main contributions of this work include identifying publishing areas for spotting fake news, analyzing linguistic idioms employed in publications and their application in future works, exploring publication trends over the years, highlighting contributions from different nations, and emphasizing the role of authors in identifying fake news. Thanh Thi Nguyen et al. [4] conducted a study on DF using DL techniques, focusing on the risks they pose to national security, democracy, and privacy. Deepfake algorithms have the ability to produce fake photos and videos that are indistinguishable from real ones, emphasizing the need for technology to automatically identify and evaluate the integrity of digital visual content. The study examined the techniques used to create DF and reviewed existing approaches for detecting them. The researchers discussed the challenges, research directions, and developments related to deepfake technologies. Their work provides a comprehensive overview of deepfake approaches and supports the development of new and more reliable methods for coping with increasingly difficult-to-detect deepfakes.

Sumaiya Thaseen Ikram et al. [8] introduced an improved DL technique for the detection of deep fakes. The study addressed the current challenges in deep fakes and DL and proposed a strategy to overcome these challenges. With the advancement of AI technology, fake videos and photos can be created, leaving behind subtle signs of manipulation. These fraudulent videos can be used in various unethical ways to intimidate, deceive, or threaten others. The study focused on DF, an AI-based method for creating synthetic versions of human photographs, which involves fusing and overlaying pre-existing videos onto original ones. The researchers developed a system that utilizes "a hybrid CNN consisting of Inception, ResNet v2, and Xception to extract frame-level features. They conducted experimental analysis using the DFDC deep fake detection challenge dataset on Kaggle, optimizing the deep learning-based techniques to improve accuracy and training time. The results showed high precision,recall,F1-score, and support, indicating the effectiveness of the proposed technique in detecting DF".

In summary, the studies discussed here highlight the ongoing efforts to detect and address DF using deep learning techniques. The research focuses on both audio-visual DF, as well as DF in news articles. These studies contribute to the development of novel architectures and algorithms, explore research trends, identify challenges, and propose strategies to improve the detection and identification of DF. By leveraging DL and related methodologies, researchers aim to mitigate the risks associated with DF and promote a safer digital environment.

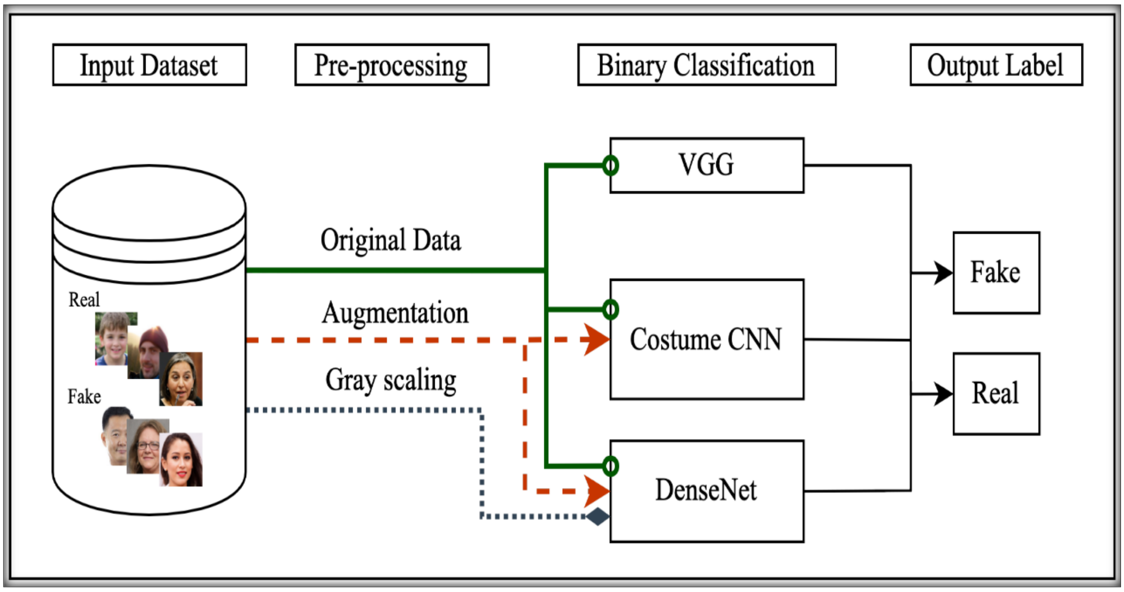

Figure 4. Sumaiya Thaseen Ikram et al. [8] used methodology

Alex Munyole Luvembe et al. [9] conducted a study on DF detection in news using DL techniques. The researchers addressed the need to automatically detect fake news to minimize its negative impact on society. While there are many methods for spotting misleading information, many of them just take data-oriented text characteristics into account, ignoring the significance of dual emotion characteristics, such as publishing emotions and social feelings, leading to lesser accuracy. The study suggested using dual emotion features for spotting fake news as a solution to this problem. They introduced "a Deep Normalization Attention-based technique for improving the extraction of dual emotion characteristics as well as an Adaptive Genetic Weight Update-RF (AGWu-RF) for classification. BiGRU was added into the deep standardized attention-based approach to enhance feature value and address gradient explosion problems brought on by far-reaching context information. Additionally, to obtain optimized parameters that increase detection accuracy, the genetic load for the model has been revised and modified to the RF classifier. Proposed model outperformed baseline techniques on common benchmark criteria, surpassing new techniques in terms of accuracy by 5%, 11%, and 14% across three real-world datasets. This highlights the significance of incorporating dual emotion capabilities and optimizations" in enhancing the identification of fake news. Ruben Tolosana et al. [10] introduced a new method for DF detection in 2023. The study focused on media forensics, which has gained attention due to the growing concerns surrounding DF. With advancements in visual techniques, distinguishing fake videos from real ones has become nearly impossible. The researchers conducted a thorough analysis of both the first and second generations of DF, utilizing two different approaches in their experimental framework. The conventional method, commonly used in the literature, involved selecting the entire face as input for the fake detection system. The cutting-edge method, on the other hand, focused on "selecting specific facial regions as input. Fusion techniques were applied to the selected face regions and three advanced fake detection systems (Xception, Capsule_Network, and DSP-FWA) to enhance the resilience of the detectors. The study also explored intra- and inter-database scenarios to strengthen the detectors against unseen attacks. The experiments revealed excellent results using facial regions and fusion techniques, achieving fake detection results above 99% AUC for the UADFV, FaceForensics++, and Celeb-DF v2 databases". The study emphasized the importance of analyzing inter-database scenarios to enhance the detectors' capabilities against previously unseen attacks. In [11], a deep fake-based survey was conducted to explore DL techniques. Momina Masood et al. [12] constructed a study that discussed the state-of-the-art challenges in DF detection using video, images, and audio datasets. The researchers emphasized the alarming trend of deep fakes, which are manufactured media intended to mislead, disseminate misinformation, commit fraud, and interfere with governments. The paper provided a comprehensive review and analysis of available tools and ML-based methods for detecting DF generation. It also discussed the techniques used to detect manipulations in both audio and video. Unlike previous surveys that mainly focused on deep fake videos and images, this study delved into manipulation techniques, public datasets, benchmarks, and open challenges. The work aimed to help researchers understand the production and detection of DF, their limitations, and potential directions for future research.

Overall, these studies contribute to the field of DF detection by proposing novel techniques, exploring the significance of dual emotion features, analyzing different approaches and fusion techniques, highlighting the challenges and limitations, and providing directions for future research. By leveraging DL and innovative methodologies, researchers aim to improve the identification and mitigation of DF media, promoting a safer and more trustworthy digital environment.

3. Proposed DF-DL Model

The DF-DL system that is being suggested starts by choosing a video dataset from a variety of freely accessible web sources, including YouTube, Amazon, and security cameras. In order to enhance dataset quality and prepare it for the deployment of deep learning models, preparation steps for the video footage (such as scaling, normalisation, and framing) are applied. The next step after removing frames or images from films is to annotate the two distinct classes (Fake and Real).

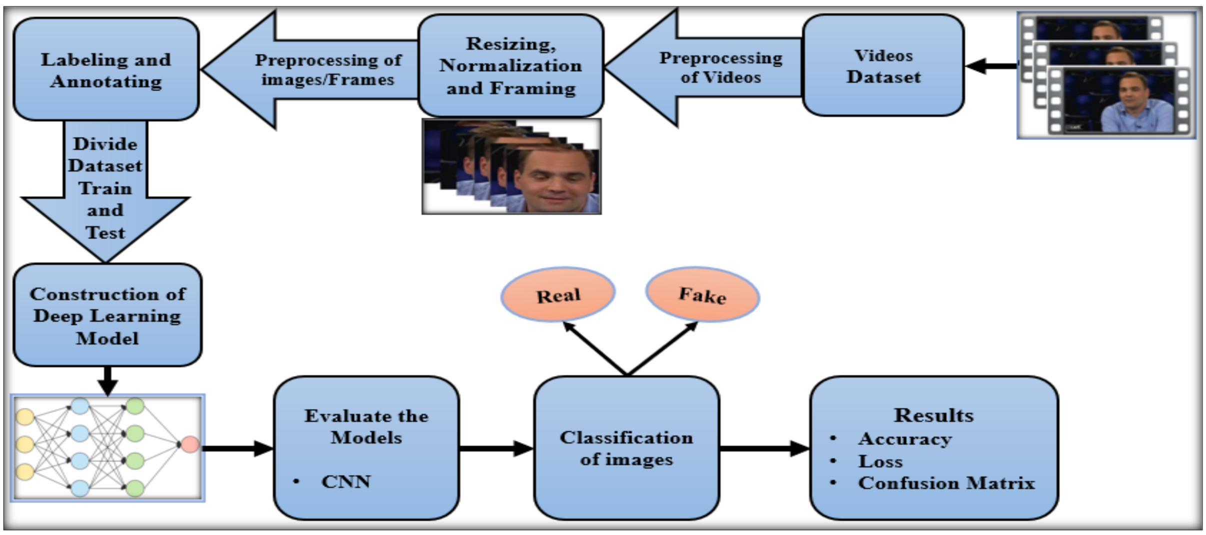

Figure 5. Proposed methodology.

In order to train and test the DL model, the pictures dataset is further segmented into training and testing datasets. Fourth, a CNN DL model. Finally, the results of the model are compared in terms of accuracy, loss and confusion matrix. Finally, the accuracy, loss, and confusion matrices of the model's output are compared.

The chosen dataset includes 402 photos from the Real and Fake classes. The videos are compiled from a variety of online channels, including Facebook, Amazon, and YouTube.

Table 1. Division of videos

Data	No. of Videos
Train	300
Test	102
Total	402

102 videos make up the testing dataset, whereas 300 videos form part of the training dataset.

Table 2. Training Dataset

Class	No. of Videos
Real	163
Fake	137
Total	300

The number of instances in the training dataset for each class are displayed in Table 2. In the training dataset, there are 137 videos from the Fake class and 163 videos from the Real class. The testing dataset's instance or video count for each class is displayed in Table 3. The training dataset includes 38 films from the Fake class while the testing dataset includes 64 videos from the Real class.

Table 3. Testing Videos

Class	No. of Videos
Real	64
Fake	38
Total	102

Convolution, max_pooling, flatten, and dense layers are among the four different types of layers in our suggested CNN model.

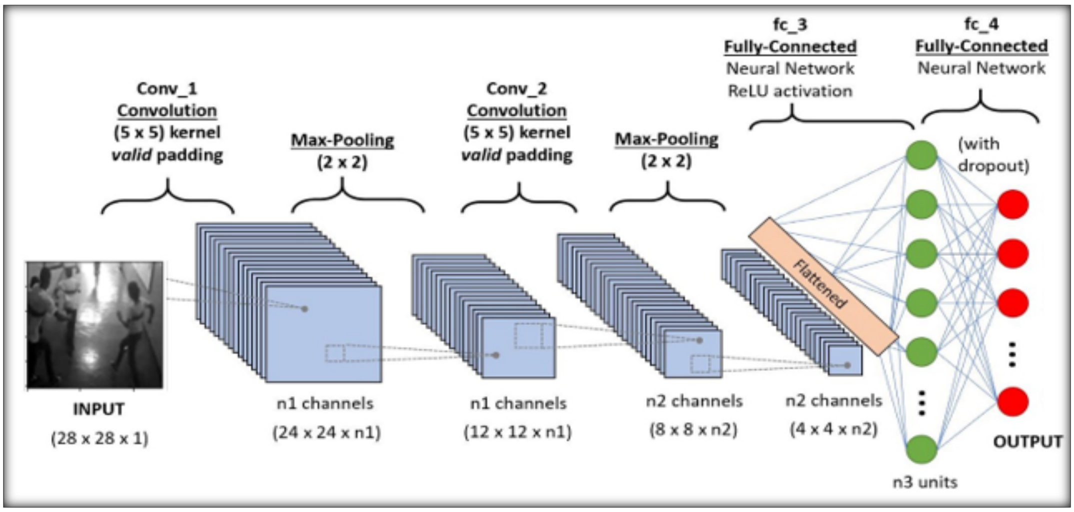

Figure 6. CNN Architecture

In machine learning, a CNN model is a class of DL models specifically designed for computer vision tasks, where the input is typically an image or a collection of images. CNNs have transformed computer vision and are now a key component of many cutting-edge methods for tasks including image classification, object identification, picture segmentation, and more. These deep learning models were created with the explicit purpose of extracting pertinent information from input photographs and making predictions based on those features. A CNN model consists of multiple layers that work together to process and understand visual data. One of the key components is the convolutional layer, which applies learnable filters, also known as kernels, to the input data. Through a convolution operation, the layer performs element-wise multiplication between the filter weights and a local region of the input, followed by summation. This process captures spatial patterns and features present in the input data, enabling the network to learn meaningful representations. To further enhance the performance of CNNs, activation functions are introduced after the convolutional layers. Activation functions "introduce non-linearity into the model, allowing it to capture complex relationships between the input and output. Common activation functions include sigmoid, tanh, ReLU (Rectified Linear Unit)", and softmax, which are chosen based on the nature of the task and the desired output requirements.

Pooling layers are another crucial element in CNNs. A popular type of pooling layer is the max pooling layer, which performs downsampling on the feature maps generated by the convolutional layers. It divides the input into non-overlapping rectangular regions, called pooling windows or pooling kernels, and independently selects the maximum value within each window. By discarding the remaining values, the layer reduces the spatial dimensions of the input, resulting in smaller feature maps. This downsampling operation helps in reducing the computational complexity of the model and providing translation invariance to small spatial variations.

In a CNN model, after several convolutional and pooling layers, the data is typically passed through fully connected layers. Before feeding the data into these layers, a flatten layer is often used to reshape the multidimensional output from the previous layers into a one-dimensional vector. The flatten layer collapses all the dimensions except the batch dimension, producing a linearized representation of the data. This transformation allows the subsequent fully connected layers to process the data effectively.

Each element in the vector represents the activation value of a specific neuron", which serves as input to the next layer in the neural network. By leveraging dense layers, deep learning models can effectively learn and capture intricate patterns in the data. To build a CNN model, the layers are typically stacked together sequentially. The input data flows through the layers in a forward pass, and the model learns the optimal weights and biases through backpropagation during the training process. The model is trained on a labeled dataset, where the input images are associated with their respective classes. The loss function measures the discrepancy between the predicted outputs and the true labels, and the optimizer adjusts the model's parameters to minimize this loss. The training process involves iterating over the training dataset for multiple epochs, gradually improving the model's performance.

In Python, popular deep learning libraries such as TensorFlow and Keras provide powerful tools for building and training CNN models. These libraries offer high-level abstractions that simplify the process of constructing the model architecture, handling the data, and performing the training and evaluation steps. In conclusion, CNNs have significantly advanced the field of computer vision by effectively capturing spatial patterns and features in images. The convolutional, pooling, flatten, and dense layers work together to extract meaningful representations from the input data and make predictions based on those features. By leveraging the power of deep learning, CNNs have achieved remarkable success in various computer vision tasks, opening up possibilities for applications in fields such as autonomous driving, medical imaging, robotics, and more.

4. Results and Discussion

On the training and testing datasets, the proposed CNN model is assessed, and the accuracy and loss of each dataset are noted. The suggested CNN's architecture is seen in the image below. The first layer (Cov2d_6) in this image contains an output with the parameters (64, 64, 32) and translates this output as an input to the following layer. The input from the first layer is transferred to the second layer, max_pooling_6, which transforms it into the shape of (32, 32, 32). The conv2d_7 layer of our suggested CNN model takes input from the conv2d_6 layer and converts it into (32, 32, 32) with 9248 parameters.

```
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 conv2d_6 (Conv2D)           (None, 64, 64, 32)        896

 max_pooling2d_6 (MaxPooling  (None, 32, 32, 32)       0
 2D)

 conv2d_7 (Conv2D)           (None, 32, 32, 32)        9248

 max_pooling2d_7 (MaxPooling  (None, 16, 16, 32)       0
 2D)

 flatten_3 (Flatten)         (None, 8192)              0

 dense_6 (Dense)             (None, 64)                524352

 dense_7 (Dense)             (None, 5)                 325

=================================================================
Total params: 534,821
Trainable params: 534,821
Non-trainable params: 0
```

Figure 7. CNN Model Configuration

The employed CNN model's architecture is depicted in Figure 7. The accuracy of the CNN model is shown in the following table.

Table 4. Used DL model Parameters

Parameter	Description
Model	Custom_CNN
No Layers	7
Epochs	50
Batch Size	64
Optimizer	Adam
Loss	Cross entropy

A custom CNN model is used for this research that 7 different layers. We use 50 epochs to evaluate and compile the model. We use batch size of 64 with adam optimizer.

4.1 Accuracy

The (Table 5) shows the training and validation accuracy of the proposed Custom_CNN model. Row one of the tables show the parameters and its description and row two shows the training accuracy of the model.

Table 5. Accuracy

Accuracy	Description

Training	0.9255
Validation	0.9486

Row third shows the validation accuracy of the proposed model. CNN model achieved training accuracy of 92.55% & validation accuracy of 94.86%.

```
Epoch 45/50
64/64 [==============================] - 51s 803ms/step - loss: 0.2583 - accuracy: 0.8947 - val_loss: 0.1507 - val_accuracy: 0.9512
Epoch 46/50
64/64 [==============================] - 51s 798ms/step - loss: 0.2298 - accuracy: 0.9157 - val_loss: 0.1389 - val_accuracy: 0.9486
Epoch 47/50
64/64 [==============================] - 54s 850ms/step - loss: 0.2426 - accuracy: 0.8966 - val_loss: 0.1476 - val_accuracy: 0.9460
Epoch 48/50
64/64 [==============================] - 54s 847ms/step - loss: 0.2209 - accuracy: 0.9128 - val_loss: 0.1185 - val_accuracy: 0.9589
Epoch 49/50
64/64 [==============================] - 50s 780ms/step - loss: 0.2286 - accuracy: 0.9147 - val_loss: 0.1360 - val_accuracy: 0.9589
Epoch 50/50
64/64 [==============================] - 54s 852ms/step - loss: 0.2027 - accuracy: 0.9255 - val_loss: 0.1318 - val_accuracy: 0.9486
```

Figure 8. Epochs of the proposed model

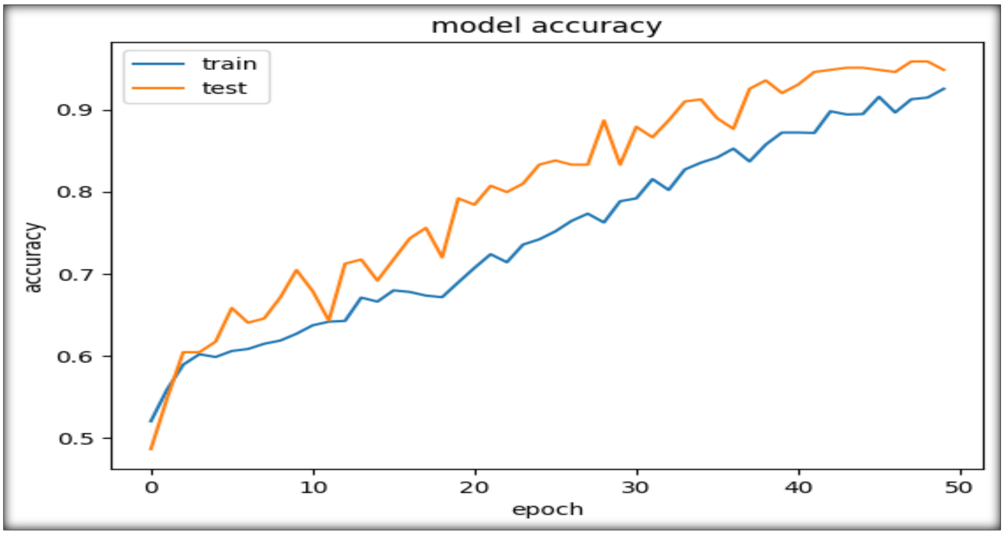

Figure 9. Model Accuracy

Figure 8 displays a screenshot of the model's most recent six epochs along with the model's training accuracy, validation accuracy, training loss, and validation loss. The graph in (Figure 9) elaborates on the accuracy of our suggested model. The graph's X-axis lists the number of epochs, while its Y-axis lists the accuracy percentage. In the graph, the blue line reflects the proposed model's training accuracy and the orange line its validation/testing accuracy. The graph demonstrates that as the number of epochs rises, accuracy in both training and testing eventually improves. The precision varies between the first and last epochs, with the last epoch having higher accuracy.

4.2 Loss

The (Table 6) shows the training and validation loss of the proposed Custom_CNN model. Row one of the tables show the parameters and its description and row two shows the training Loss of the model.

Table 6. Loss

Parameter	Description

Training	0.2027
Validation	0.1318

Row third shows the validation Loss of the proposed model. CNN model gives training loss of 20.27% and validation Loss of 13.18%. The Loss of our proposed model is displayed in the graph of (Figure 10). The graph contains number of epochs in X-axis and the percentage of loss is given in the Y-axis of the graph. Blue line represents the training loss and orange line in the graph represent the validation loss of the proposed model. The graph shows that loss of both training and testing is gradually decreases when the number of epochs is increases. At the first epoch the loss is maximum and at the last epoch the accuracy is lowest.

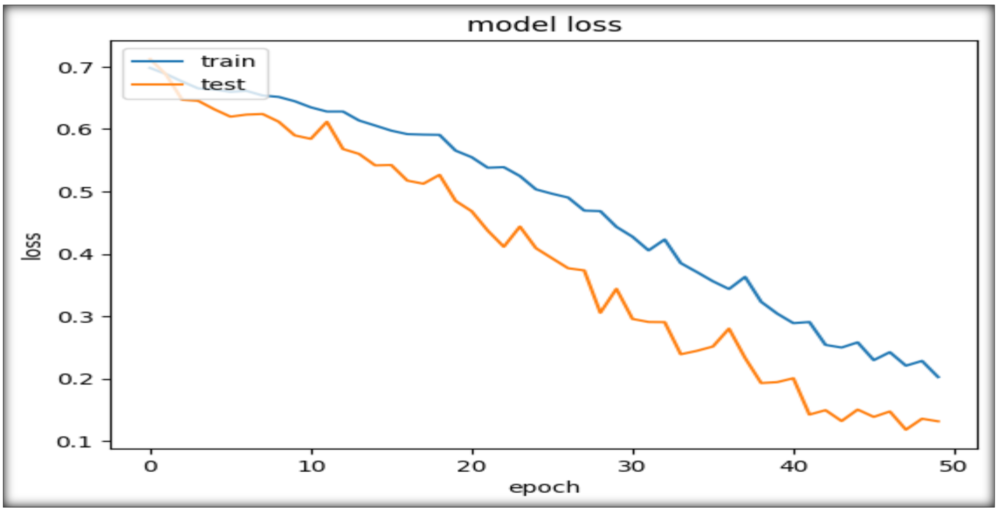

Figure 10. Model Loss

4.3 Confusion Matrix

The confusion matrix results (Recall, F1-Score and Recall) of the model is given the below table with each class:

Table 7. Confusion Matrix results

Label	Recall	F1-Score	Recall
Real	56	53	51
Fake	44	46	49
Average	50	50	50

The proposed model achieved recall of 56 percent, 44 percent and 50 percent of class real, fake and average. Precision of 51 percent is achieved by the class real and 49 percent by the class fake. The average precision of both the class is calculated as 50 percent. The f1-score of models is maximum at class real with accuracy of 53 percent and minimum at class fake with accuracy of 46 percent. (Figure 11) shows the confusion matrix graph from the prediction proposed model X-axis shows the predicted labels that predicted by the proposed model and Y-axis contains the true labels of the images. There are 389 images are taken for the testing of model. The propose model predicted 200 images as 0 (Real) in which 111 images are predicted as true and remaining 89 images are predicted wrong. 189 images from class 1 (Fake) are predicted in which 84 images are predicted accurately and 105 images are predicted wrongly.

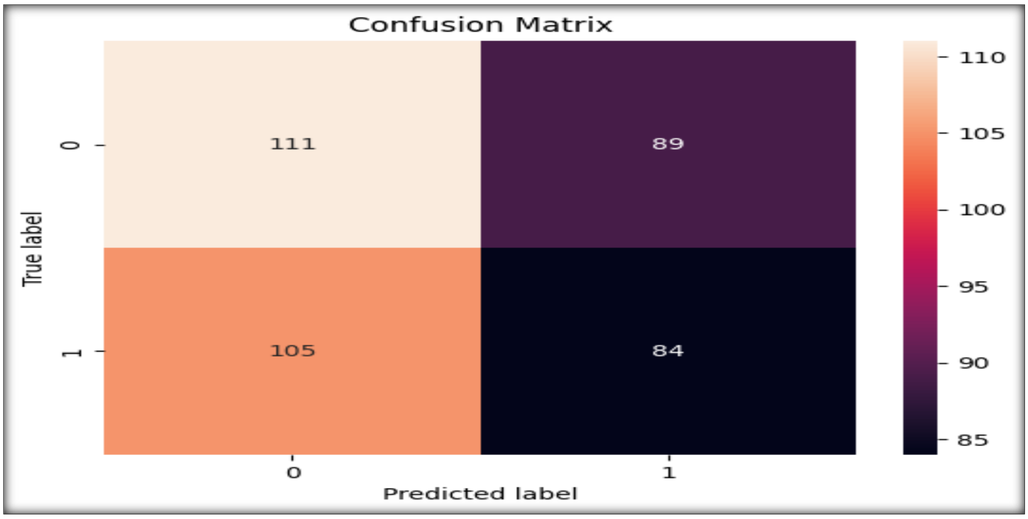

Figure 11. Confusion Matrix

4.4 Comparison

Comparison of proposed model with other algorithems are given in the below section:

Table 8. Model Comparison

Models	Training Accuracy	Validation Accuracy
Proposed	93%	95%
ResNet	89%	91%
Inception V3	76%	90%
ViT	90%	85%

Our proposed model performed well on the processed dataset as compared to other models. ResNet model achieved training&validation accuracy of 89 percent and 91 percent. The Inception model is training poor with accuracy only 76 percent and validation accuracy of 90 percent.

5. Conclusion

Deep fake technology has recently advanced significantly, enabling the creation of highly realistic fake audio, video, and image media. These materials pose significant challenges, as they can lead to impersonation, dissemination of false information, and potential national security risks that compromise authentication. To address these issues, an arms race has emerged between deep fake creators and detectors, prompting the development of various deep fake detection algorithms. However, these detectors often prove unreliable and frequently fail to identify deep fakes.

This study proposed a deep learning method to mitigate the difficulties associated with identifying deep fakes in video datasets. In this study, a neural network model was proposed for deep fake detection in videos. Initially, a video dataset was selected from the well-known Kaggle dataset repository. The dataset was then augmented into two classes: real videos and fake videos, followed by its division into training and testing subsets. Next, the dataset underwent preprocessing, involving face extraction, region of interest selection, and frames extraction to discern between real and fake videos. Subsequently, a neural network was applied to the processed dataset, and the model's performance was evaluated by calculating its accuracy. Finally, a comparison was conducted between the proposed model and state-of-the-art models such as ResNets, Inception V3, and vision transformers. The comparison demonstrated that our proposed model performed favorably on the processed dataset compared to other models, achieving an accuracy of 94.86 percent.

References

1. Tolosana, R., Vera-Rodriguez, R., Fierrez, J., Morales, A., & Ortega-Garcia, J. (2020). Deepfakes and beyond: A survey of face manipulation and fake detection. Information Fusion, 64, 131-148.
2. Ilyas, H., Javed, A., & Malik, K. M. (2023). AVFakeNet: A unified end-to-end Dense Swin Transformer deep learning model for audio–visual deepfakes detection. Applied Soft Computing, 136, 110124.
3. Dhiman, P., Kaur, A., Iwendi, C., & Mohan, S. K. (2023). A scientometric analysis of deep learning approaches for detecting fake news. Electronics, 12(4), 948.
4. Nguyen, T. T., Nguyen, C. M., Nguyen, D. T., Nguyen, D. T., & Nahavandi, S. (2019). Deep learning for deepfakes creation and detection. arXiv preprint arXiv:1909.11573, 1(2), 2.
5. Mirsky, Y., & Lee, W. (2021). The creation and detection of deepfakes: A survey. ACM Computing Surveys (CSUR), 54(1), 1-41.
6. Korshunov, P., & Marcel, S. (2018). Deepfakes: a new threat to face recognition? assessment and detection. arXiv preprint arXiv:1812.08685.
7. Akhtar, Z. (2023). Deepfakes Generation and Detection: A Short Survey. Journal of Imaging, 9(1), 18.
8. Ikram, S. T., Chambial, S., & Sood, D. (2023). A performance enhancement of deepfake video detection through the use of a hybrid CNN Deep learning model. International journal of electrical and computer engineering systems, 14(2), 169-178.
9. Luvembe, A. M., Li, W., Li, S., Liu, F., & Xu, G. (2023). Dual emotion based fake news detection: A deep attention-weight update approach. Information Processing & Management, 60(4), 103354.
10. Tolosana, R., Romero-Tapiador, S., Vera-Rodriguez, R., Gonzalez-Sosa, E., & Fierrez, J. (2022). DeepFakes detection across generations: Analysis of facial regions, fusion, and performance evaluation. Engineering Applications of Artificial Intelligence, 110, 104673.
11. Rahman, A., Islam, M. M., Moon, M. J., Tasnim, T., Siddique, N., Shahiduzzaman, M., & Ahmed, S. (2022). A qualitative survey on deep learning based deep fake video creation and detection method. Aust. J. Eng. Innov. Technol, 4(1), 13-26.
12. Masood, M., Nawaz, M., Malik, K. M., Javed, A., Irtaza, A., & Malik, H. (2022). Deepfakes Generation and Detection: State-of-the-art, open challenges, countermeasures, and way forward. Applied Intelligence, 1-53.
13. Xu, Y., Raja, K., & Pedersen, M. (2022). Supervised contrastive learning for generalizable and explainable deepfakes detection. In Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision (pp. 379-389).
14. Shahzad, H. F., Rustam, F., Flores, E. S., Luís Vidal Mazón, J., de la Torre Diez, I., & Ashraf, I. (2022). A Review of Image Processing Techniques for Deepfakes. Sensors, 22(12), 4556.
15. Ilyas, H., Javed, A., & Malik, K. M. (2023). AVFakeNet: A unified end-to-end Dense Swin Transformer deep learning model for audio–visual deepfakes detection. Applied Soft Computing, 136, 110124.
16. Li, W., Guo, C., Deng, Z., Liu, F., Wang, J., Guo, R., ... & Jin, Q. (2023). Coevolution modeling of group behavior and opinion based on public opinion perception. Knowledge-Based Systems, 110547.
17. Liu, Z., Qi, X., & Torr, P. H. (2020). Global texture enhancement for fake face detection in the wild. In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition (pp. 8060-8069).
18. Kim, Y., Na, J., Yoon, S., & Yi, J. (2009). Masked fake face detection using radiance measurements. JOSA A, 26(4), 760-766.
19. Maqbool, M. S., Hanif, I., Iqbal, S., Basit, A., & Shabbir, A. (2023). Optimized Feature Extraction and Cross-Lingual Text Reuse Detection using Ensemble Machine Learning Models. Journal of Computing & Biomedical Informatics, 5(01), 26-40.
20. Fazal, U., Khan, M., Maqbool, M. S., Bibi, H., & Nazeer, R. (2023). Sentiment Analysis of Omicron Tweets by using Machine Learning Models.
21. Hasnain, M. A., Ali, S., Malik, H., Irfan, M., & Maqbool, M. S. (2023). Deep Learning-Based Classification of Dental Disease Using X-Rays. Journal of Computing & Biomedical Informatics, 5(01), 82-95.
22. Basit, A., Hanif, I., Maqbool, M. S., Qayyum, W., Hasnain, M. A., & Nazeer, R. (2023). Cross-Lingual Information Retrieval in a Hybrid Query Model for Optimality. Journal of Computing & Biomedical Informatics, 5(01), 130-141.
23. Wang, C., & Deng, W. (2021). Representative forgery mining for fake face detection. In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition (pp. 14923-14932).
24. Guo, Z., Yang, G., Chen, J., & Sun, X. (2021). Fake face detection via adaptive manipulation traces extraction network. Computer Vision and Image Understanding, 204, 103170.
25. Do, N. T., Na, I. S., & Kim, S. H. (2018). Forensics face detection from GANs using convolutional neural network. ISITC, 2018, 376-379.
26. Mansourifar, H., & Shi, W. (2020). One-shot gan generated fake face detection. arXiv preprint arXiv:2003.12244.
27. Dang, H., Liu, F., Stehouwer, J., Liu, X., & Jain, A. K. (2020). On the detection of digital face manipulation. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern recognition (pp. 5781-5790).
28. Pan, Z., Ren, Y., & Zhang, X. (2021). Low-complexity fake face detection based on forensic similarity. Multimedia Systems, 27, 353-361.
29. Liu, Z., Qi, X., & Torr, P. H. (2020). Global texture enhancement for fake face detection in the wild. In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition (pp. 8060-8069).
30. Wang, C., & Deng, W. (2021). Representative forgery mining for fake face detection. In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition (pp. 14923-14932).