Artificial Illusion: The Art of Deepfake Technology

Written by:

Jawaher Mohammed Alkhamis (g202403980)

Dhoha Ahmed Almubayedh (g202403920)

Supervised by:

Muzammil Behzad

KING FAHD UNIVERSITY OF PETROLEUM AND MINERALS

MX - PROGRAM

26 April 2025

Table of Contents

Abstract

In this report, we trained a 3D convolutional model and retrained the original model on different datasets. We also improved the model's robustness by using the AdamW optimizer with weight decay. After that, we tested the models on datasets of various sizes and tracked their performance by plotting the training and validation curves. Overall, the results showed slight improvements.

Introduction

Artificial Intelligence (AI) has become one of the most widely discussed topics across various fields in recent years, sparking both curiosity and concern due to its rapid advancements. Among AI-driven technologies, deepfake stands out as a powerful yet controversial tool. It enables the creation of synthetic media that can convincingly impersonate real individuals, objects, or events, often with striking realism.

Due to its ability to generate highly realistic fake content with minimal effort, deepfake technology has significantly impacted the media industry. Unfortunately, its accessibility also makes it a tool for malicious purposes, such as spreading misinformation, manipulating public opinion, and committing fraud. Scammers, for instance, can easily create faked media to impersonate real individuals, deceiving the public and spreading false narratives.

Considering these concerns, we've searched for a way to detect these falsifications and found a suitable proposed methodology in the research paper *"MesoNet: A Compact Facial Video Forgery Detection Network"* [1].

Problem Statement

As deepfake generator continue to spread on various fields the detection models are few and less compatible, though there do exist some strong models that have strengthen their root as best detectors for deepfake data such as Deepfake Detection Challenge (DFDC)[2]. To address this, we aim to improve the performance of the MesoNet model, which has shown promise in detecting deepfakes and Face2Face manipulations.

Objectives

This is a report that focuses on enhancing the following parts in the MesoNet:

1. Improve the model to be using 3D convolution for video detection
2. Enhance the model to be using AdamW optimization
3. Enhance the output of the model to be understandable and concise
4. Letting the model be more feasible to general data
5. Finetune the training model and the pipelines
6. Improve the model by adding an Ensimble learning model

Scope Of Study

This report focuses on the analysis and enhancement of the MesoNet deepfake detection model. It includes retaining the model on new supportive added models and datasets to be more adaptable to generalization, and finetune the pipeline alongside the output to be meaningful and clear.

Literature Review

Recent Methods and Application

The rapid rise of deepfake techniques resulted in an increasing need for reliable detection methods. Conducted studies proposed several techniques to identify manipulated facial videos to ensure the authenticity of the digital content.

One of the earliest studies in this domain is "FaceForensics++: Learning to Detect Manipulated Facial Images.", which presented a large-scale dataset containing both real and deepfake videos, making it a valuable benchmark for training detection models. The dataset discussed in this study covered different manipulation techniques such as Face2Face, FaceSwap and DeepFakes. By leveraging convolutional neural networks (CNNs), this work illustrated how deepfake detection models can efficiently identify artifacts indicates the manipulation activities of real videos. Nevertheless, while models based on CNN performed well on controlled datasets, they often struggled to generalize to new, unseen deepfake techniques. This limitation highlighted the need for more adaptable detection strategies.

Another study presented by Zhang et al, in "Exploring Temporal Coherence for More General Video Face Forgery Detection", which explored the detection of deepfake video through motion inconsistencies in deepfake videos. This research proposed the use of Fully Temporal Convolution Network (FTCN), which captures inconsistencies across multiple frames. By analyzing the motion flow and tracking facial dynamics over time, which significantly improved detection accuracy, mainly for deepfakes that attempt to maintain high frame-by-frame consistency. However, while this approach was highly effective, it introduced computational overhead, making real-time implementation more difficult.

In contrast, Prajapati and Pollett in their study titled "MRI-GAN: A Generalized Approach to Detect DeepFakes using Perceptual Image Assessment.", they gave an innovative approach to detect deepfake videos by applying Generative Adversarial Networks (GANs) to evaluate perceptual quality inconsistencies in deepfake videos, unlike traditional CNN-based detection methods. This study focused on texture, resolution, and structural distortions, which made MRI-GAN provide a good generalization rate across different deepfake techniques without being restricted to dataset-specific artifacts. While MRI-GAN achieved state-of-the-art accuracy, its high computational cost posed a challenge for real-time detection, making it impractical for large-scale deployment.

Recently, Lanzino et al, in "Faster Than Lies: Real-time Deepfake Detection Using Binary Neural Networks" addressed the issue of computational efficiency in deepfake detection

techniques by introducing Binary Neural Networks (BNNs) to reduce processing time while maintaining strong detection performance. Contrasting to traditional CNN-based models, BNNs allow for faster inference and deployment on low-resource devices, making them particularly suitable for real-time applications such as live-streamed content monitoring. However, while BNNs improve efficiency, they may experience minor accuracy reductions compared to more complex deep learning models.

Gaps and Limitation

Despite rapid progress, state-of-the-art video deep-fake detectors still share weaknesses. Most of deepfake detectors depend on frame-level processing at a time and therefore ignore temporal artifacts that show up on successive frames. Also, the training is based on specific deepfake techniques that keeps changing with public tool-chains, leaving it difficult for the models when they are confronted with emerging, unseen forgery techniques. In addition to that, they often depend on a single face-cropping pipeline, so a minor tracking failure can erase all downstream evidence. These limitations collectively underscore the need for enhanced architectures for deepfake detections in different medias.

Proposed Model

Proposed Methodology

The proposed techniques to enhance the MesoNet model involve several points that are needed to be achieves as it's stated in the objectives. The process will be performed as follows:

1. **Data collection:**
   The model will be trained and evaluated on general – self gathered for videos- and deepfake datasets, which provide labeled video content for detection.
2. **Data Preprocessing:**
   The frames will be extracted from the videos and images and then get detected by face detection function. Lastly the frames will be resized, normalized and aligned to be able to use in training and evaluation.
3. **Architectural design and modifications:**
   The model will be modified by adding a 3D convolution layer for a better capturing face from the videos. Level up the optimizer to be AdamW to improve efficiency and accuracy especially for large datasets. refine the output of the model to be more concise, meaningful and actionable.
4. **Training**
   The model will be train on 16 batches, 60 epoch, 1e-5 startup decaying weight, 0.001 starting learning rate and 10 frames for videos. The model will avoide early stopping instead will have a learning rate scheduler and call back to decrease the learning rate and decaying weight if there was no improvement within 5 epochs.
5. **Evaluation**
   The performance will be evaluated by using accuracy, classification report, loss, and confusion matrix and real world effictivness.

6. **Tools and frameworks**
   The model is going to be implemented by using Keras with Tensorflow backend. The training, evaluation, test will be implemented using Google Colab, Kaggle, and Visual Studio code. Additional used tools and libraries are face_recognition, imageio, sikitlearn, numpy, os and matplotlib.

Proposed Model Challenges

To refine and enhance the MesoNet model, we encountered a series of significant challenges, including:

1. **Dataset Availability:**
   The original dataset link provided by the authors was no longer functional. We contacted the authors and, in the meantime, searched for alternative datasets. Eventually, we received a dataset response, but it only included image-based data for training and validation. As a result, we curated our own dataset by combining various types of fake videos from the *FaceForensics++* [3] dataset to create a semi-general dataset. Additionally, we gathered an image dataset to test the model's performance on unseen data.
2. **Incomplete Training Model**
   The baseline project only included the testing code, classifiers, and pipelines—there was no training code. This led us on a "treasure hunt" to find implementations of the training pipeline. We eventually discovered a community-contributed training script referencing the original paper, which was a major breakthrough. However, for the 3D convolution version (Meso3D), there was still no training implementation. We had to create, train, and test this model from scratch, which proved to be one of the most technically demanding parts of the project.
3. **Video Incompatibility**
   Although the original paper claimed video detection support, the implementation was designed primarily for static images and testing only. We had to revise and fine-tune the pipeline extensively to enable proper video input and processing, as expected from a deepfake detection system.
4. **Unused and Misleading Models**
   One of the models included in the repository was not referenced or used anywhere in the codebase. After investigating its structure and purpose, we concluded it was either deprecated or mistakenly added, and we chose to omit it.
5. **Unclear Output**
   The original model output was difficult to interpret. It mostly displayed time elapsed and vague accuracy predictions, making it unclear whether the model was actually functioning as expected. We improved the output format to make the results more interpretable and informative.
6. **High Computational Load**

Video preprocessing was time-consuming. Even for a small dataset, it could take over two hours, and for a mid-sized dataset, it exceeded eight hours on a standard GPU setup.

7. **Dependency and Installation Issues**
   Many dependencies required manual installation via terminal. Notably, the face_recognition library required CMake and additional configurations to function properly, which was not initially documented.
8. **Looping Preprocessing Error**
   A persistent error occurred where videos were being preprocessed repeatedly at each epoch instead of being cached, leading to redundant computation and extended training time.
7. **Run Time complixity**
   Running time for a self-gathered dataset that consist of in total 342 videos (200 for training, 88 for validating and 54 for testing) collected from different deepfake techniques to genralize the model, took approximiatly over 10 hours on GPU.

```
Preprocessing Videos:  20%|▌  | 40/200 [10:21:54<30:
Face Extraction Report:
Total number of faces found: 493
No missing data, Success!

Preprocessing Videos:  20%|▌  | 41/200 [10:32:03<29:
```

Personal Challenge

As students and full-time employees, we also faced time management difficulties, especially during Ramadan, Eid holidays, and balancing academic responsibilities with work schedules.

Proposed Loss Functions And Optimizations

To further refine the model, we replaced the Adam optimizer with AdamW, aiming to improve both accuracy and robustness by introducing better weight decay regularization.

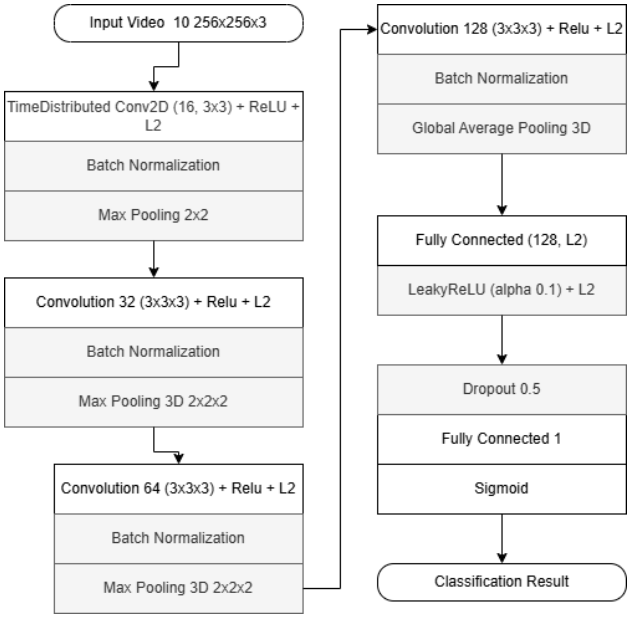
Additionally, we changed the loss function from Mean Squared Error (MSE) to Binary Cross-Entropy, which is more appropriate for binary classification tasks, helping the model better distinguish between real and fake inputs.

Model Experiments

Test Model Design and Test Dataset with Preprocessing

To thoroughly explore and validate the enhanced model as part of this project, a set of structured experiments have been involved including using the exising Meso4 and MesoInception4 with enhanced hyperparamters, and introducing 3D convolution as part of Meso model (shown below). The goal was to analyze and measure improvements against the baseline mode, specifically with the integration of 3D convolution layers shown below, use of Binary Cross Entropy loss function, AdamW optimizer, and refined output interpretation.

Model Experiments



Initially to test the integration of Binary Cross Entropy loss function instead of MSE, and AdamW optimizer instead of Adam optimizer, two distinct dataset formats have been selected: one is the official deepfake provided the author of the base research which was preprocessed image-based deepfake dataset extracted from deepfake video's frames, and the other one is video-based deepfake dataset containing .mp4 files taken from existing benchmark datasets of different types of forged videos which is FaceForensics++ dataset. FaceForensics++ covered various deepfake techniques such as FaceSwap, Face2Face, and general Deepfake manipulations. The initial target to test was on the general Deepfake manipulations videos provided by this FaceForensics++ benchmark datasets to simulate the dataset provided by the base research author but in video-based fashion.

The original dataset provided by the author contained preprocessed, labeled images intended for binary classification (Real vs. Fake). It contained general deepfake-based images, where in total there were 19457 images (cropped frames from videos divide into fake (0) / real (1) subfolders for labeling). 12353 of them for training and 7104 images for validation. To have better experiments with the updated models (Meso4 and MesoInception), the original dataset's distribution was modified to reflect 80% for training, 10% validation, and 10% for test. This structured dataset division was essential in maintaining consistency with the authors' original methodological approach, while also allowing direct comparisons to measure the effectiveness and generalizability of our proposed enhancements.

On the other hand, for the deepfake-based extracted from the FaceForensics++ dataset, a totoal of 375 videos (fake and real) were extracted from each of the 1,000 videos available in both the deepfake and original folders in FaceForensics++. These 375 videos were divided into 80% for training, 10% for validation, and 10% for test. This number of videos were extracted to make the run time manageable by the existing resources. Each video was processed by

extracting frames using the exiting pipeline with the help of face_recognition, ensuring precise facial alignment, normalization, and resizing to a consistent 256x256 image resolution.

Following preprocessing, the experiments involved training the modified models on both original and custom datasets. For Meso3D classification model, the training occurred over 60 epochs with batch sizes set to 16, starting at a learning rate of 0.001 and implementing a learning rate scheduler to reduce the learning rate and decay weights when the model showed no improvement over five consecutive epochs. Unlike traditional approaches, early stopping was avoided for comprehensive monitoring of performance variations across epochs.

For Meso4 and MesoInception4, multiple training configurations were explored, varying the number of epochs (10, 50, 60, and 100), batch sizes (16, 32, and 64), and starting learning rates (0.001, 0.005, and 0.0005…etc.) with learning rate scheduler.

Test Performance Metrics

To evaluate the model's performance, different performance metrics were used including Accuracy, Precision, Recall (Sensitivity), Specificity, F1-score, and Binary Cross-Entropy loss and included in a comprehensive classification report. Accuracy as the measure of correctness, while Precision and Recall offered into the experience an insight into the model's reliability in identifying deepfake (true positives) and authentic content (true negatives). F1-score provided a balanced harmonic mean, essential for addressing imbalanced classification scenarios.

In addition to that, the confusion matrix was used to visualize numbers related to true vs false predictions, thus simplifying the assessment of misclassification patterns.

To provide more clear results to the quantitative evaluation, different visual inspections were also performed, this includes a visual representation of Training vs. validation loss/accuracy curves, Accuracy-vs-loss scatter, Weight-decay vs LR trace, and ROC curve (val).

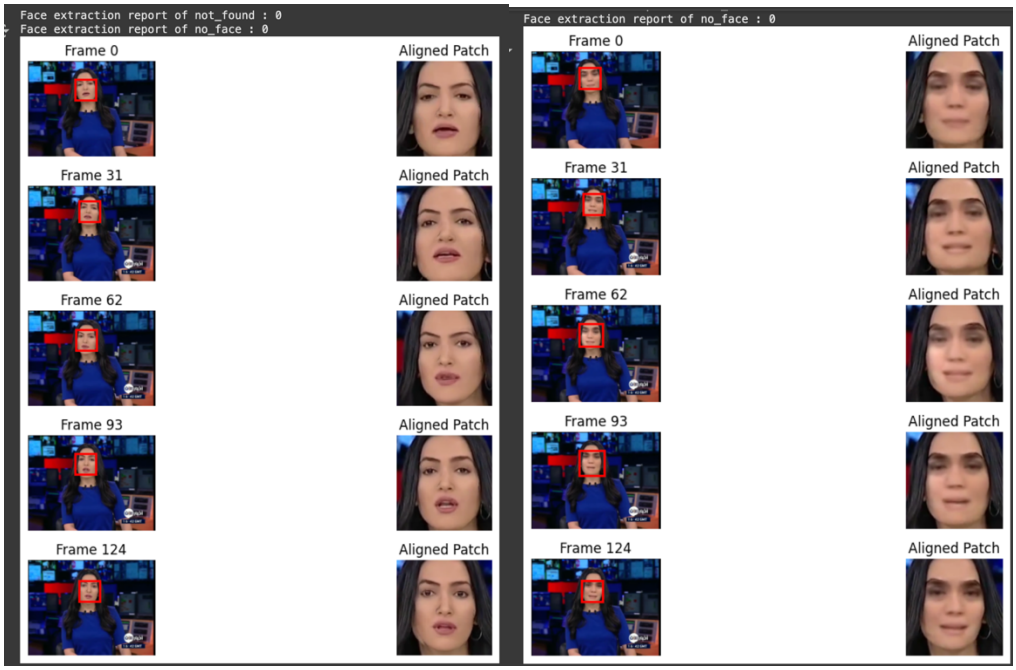
Moreover, to do additional test into the performance of extracted models, the measuring the detection rate was also included at the end of the test. The updated models have been assessed on a fixed test set of 50 deepfake videos (.mp4). The number of correctly identified deepfake videos offered a clear measure and real indicator of real-world effectiveness. This approach provided practical insight into the model’s ability to function accurately under realistic scenarios.

Model Decomposition and Results

Initially, we validated the performance of the authors' original models (Meso4 and MesoInception4) using their existing codebase provided in the official GitHub repository. The preliminary tests were conducted using the authors' predefined test set consisting of four images, achieving a perfect accuracy (100%), as indicated by the confusion matrix and classification report. These results align closely with the authors' claims of achieving over 98% accuracy for Deepfake and approximately 95% for Face2Face manipulations.

To further assess the effectiveness and robustness of the baseline preprocessing pipeline, mainly regarding face detection and frames extraction, we test the video pipeline on videos

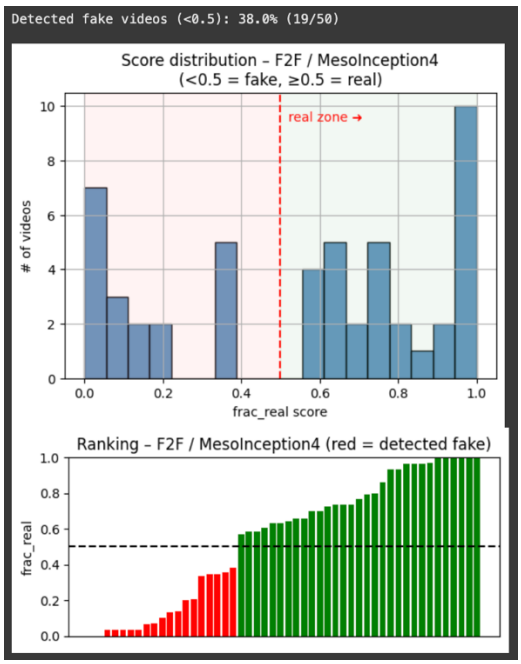
extracted from FaceForensics++ dataset (face2face and deepfake videos) as shown in the image below. This step ensured accurate evaluation of cropping accuracy, face alignment, and overall processing efficiency before integrating our proposed enhancements.

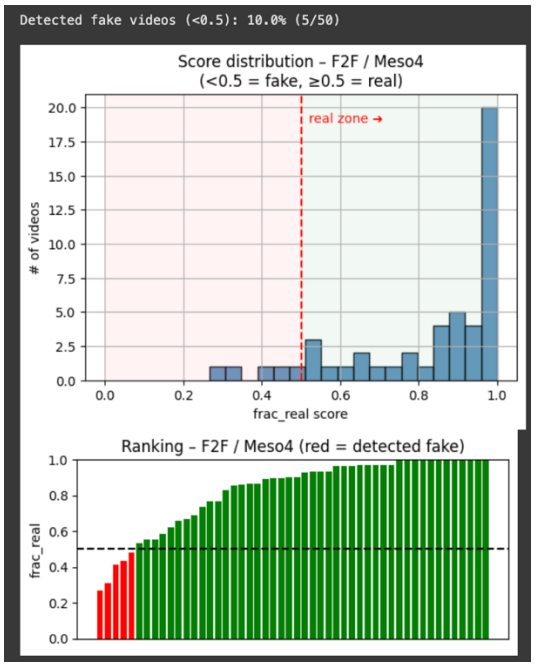


Subsequently, the baseline Meso4 and MesoInception4 models were rigorously tested against video-based datasets comprising 50 videos each from FaceForensics++ (Deepfake and Face2Face techniques). The summary of the obtained results was as follows:

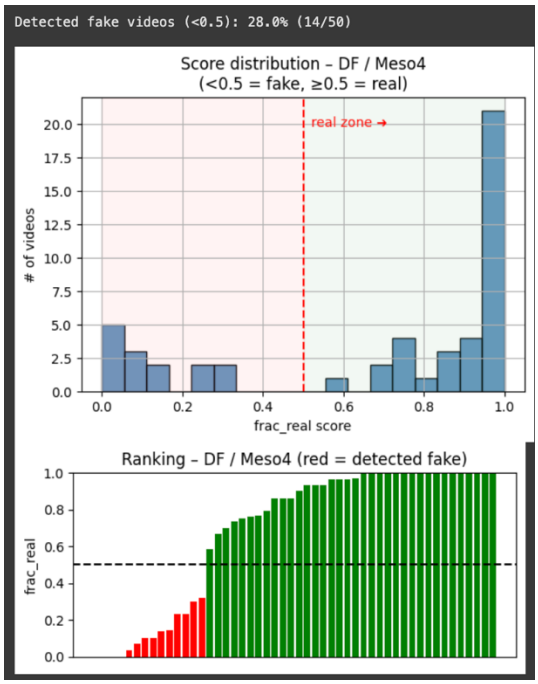
```
================  OVERALL SUMMARY  =================
set | model           | weights                 | fake%
---------------------------------------------------------
F2F | MesoInception4  | MesoInception_F2F.h5    | 38.0%
F2F | Meso4           | Meso4_F2F.h5            | 10.0%
DF  | MesoInception4  | MesoInception_DF.h5     | 12.0%
DF  | Meso4           | Meso4_DF.h5             | 28.0%
```

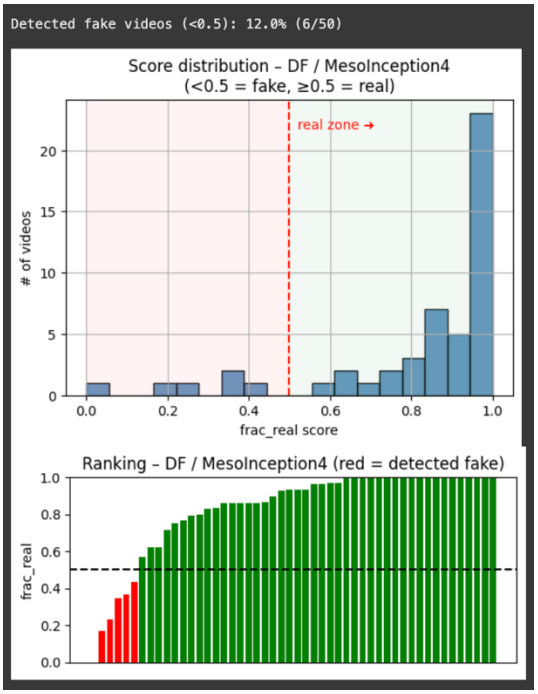
Meso4 and MesoInception4 detection distribution with Face2Face Dataset





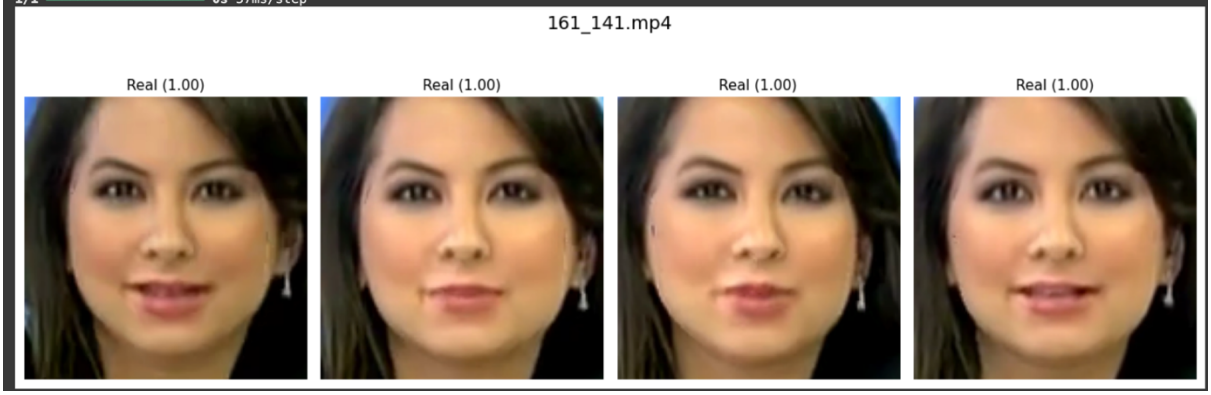
Meso4 and MesoInception4 detection distribution with DeepFake Dataset





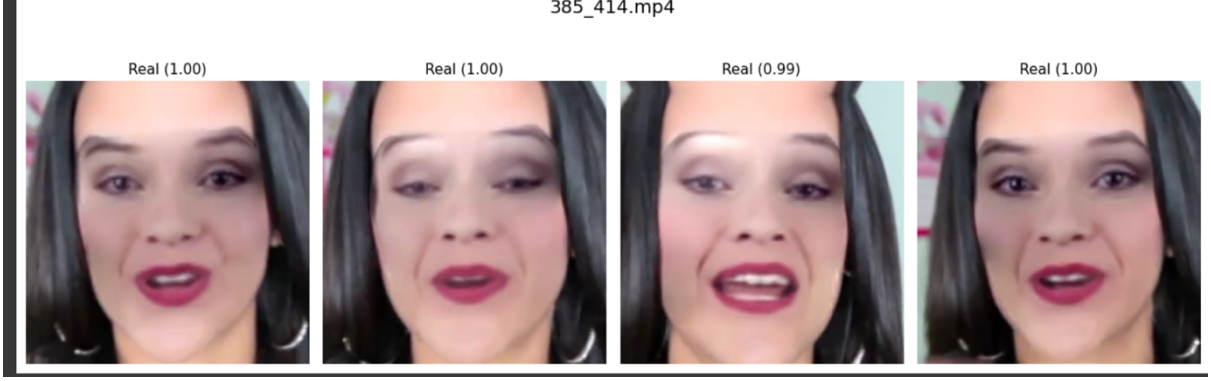
Quick look into detection rate over frames within the same video

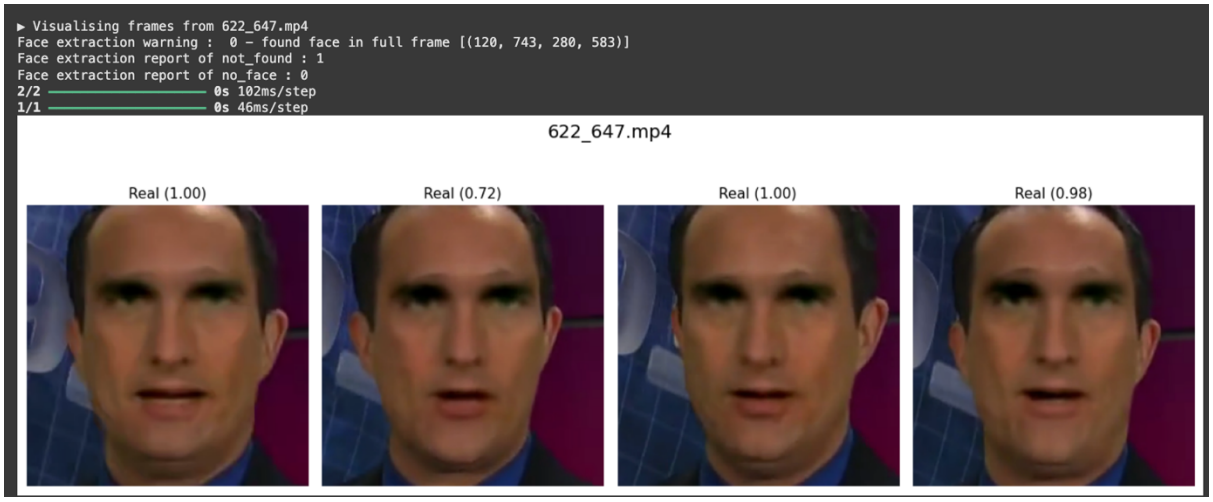
```
==========  QUICK LOOK  — F2F  ==========

▶ Visualising frames from 161_141.mp4
Face extraction warning :  0 — found face in full frame [(95, 392, 207, 280)]
Face extraction report of not_found : 1
Face extraction report of no_face : 0
2/2 ━━━━━━━━━━━━━━━━━━━━ 0s 212ms/step
1/1 ━━━━━━━━━━━━━━━━━━━━ 0s 57ms/step
```




```
▶ Visualising frames from 013_883.mp4
Face extraction warning :  0 — found face in full frame [(120, 404, 231, 293)]
Face extraction report of not_found : 1
Face extraction report of no_face : 0
2/2 ━━━━━━━━━━━━━━━━━━━━ 0s 82ms/step
1/1 ━━━━━━━━━━━━━━━━━━━━ 0s 50ms/step
```


```
==========  QUICK LOOK  — DF  ==========

▶ Visualising frames from 385_414.mp4
Face extraction warning :  0 — found face in full frame [(85, 732, 362, 454)]
Face extraction report of not_found : 1
Face extraction report of no_face : 0
2/2 ━━━━━━━━━━━━━━━━━━━━ 1s 630ms/step
1/1 ━━━━━━━━━━━━━━━━━━━━ 0s 73ms/step
```





These baseline tests indicated notable discrepancies compared to the authors' reported performance, suggesting potential limitations in model generalizability and preprocessing robustness, thus underscoring the necessity for the enhancements proposed in our research.

Training Enhanced Model and Results

Following the baseline evaluation, we developed a comprehensive training pipeline for the Meso4 and MesoInception4 models, as this functionality was initially absent from the authors' provided resources. We explored training the Meso4 and MesoInception4 models with various hyperparameters to identify optimal configurations for improved deepfake detection performance. Multiple training iterations were executed with adjustments in epochs, batch size, and learning rates, alongside utilizing advanced training techniques such as data augmentation, the AdamW optimizer, and Binary Cross-Entropy loss. For the final selected training configuration, we utilized the following initial training hyperparameters:

- Meso4: {Epochs: 60, Batch size: 64, starting learning rate: 0.001 (then with ReduceLROnPlateau scheduler), Data augmentation: Enabled}
- MesoInception4: {Epochs: 60, Batch size: 64, starting learning rate: 0.001 (then with ReduceLROnPlateau scheduler), Data augmentation: Enabled}

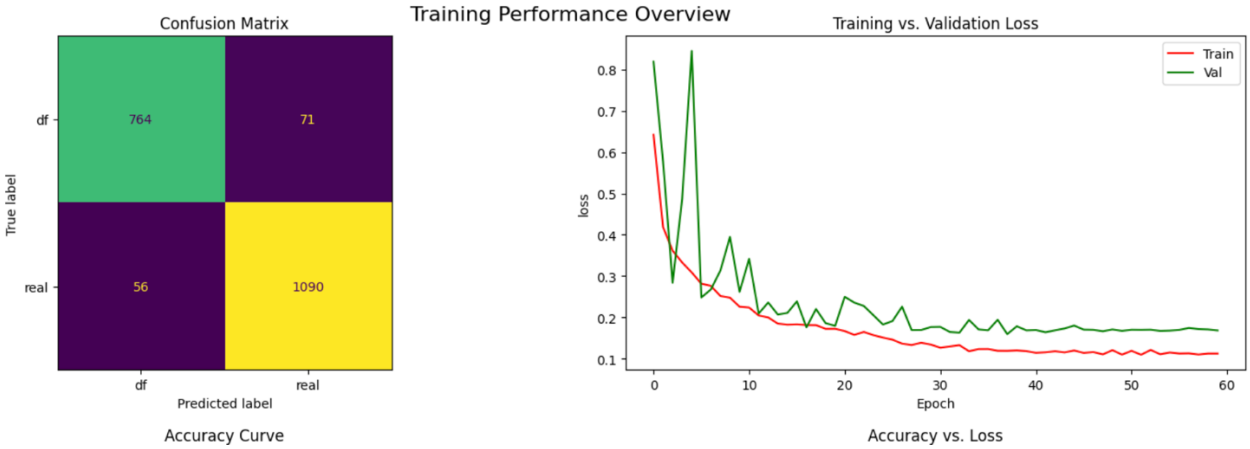
The training results demonstrated the following:

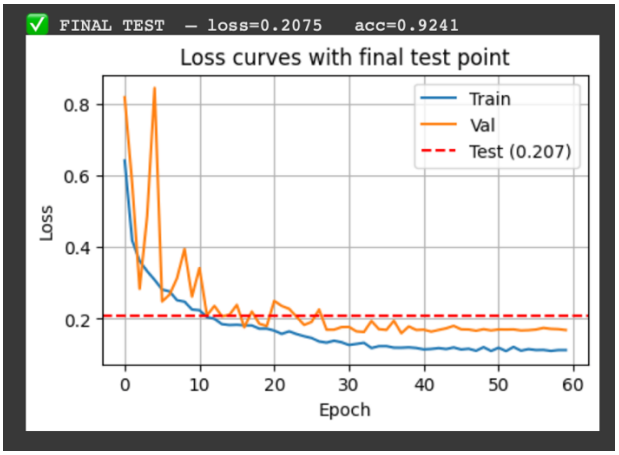
- Meso4 Training Results:
  - Best validation accuracy approximately 93.59% at epoch 20.
  - Final test accuracy was 92.41%.

Classification report (threshold = 0.50)

	precision	recall	f1-score	support
df	0.932	0.915	0.923	835
real	0.939	0.951	0.945	1146
accuracy			0.936	1981
macro avg	0.935	0.933	0.934	1981
weighted avg	0.936	0.936	0.936	1981

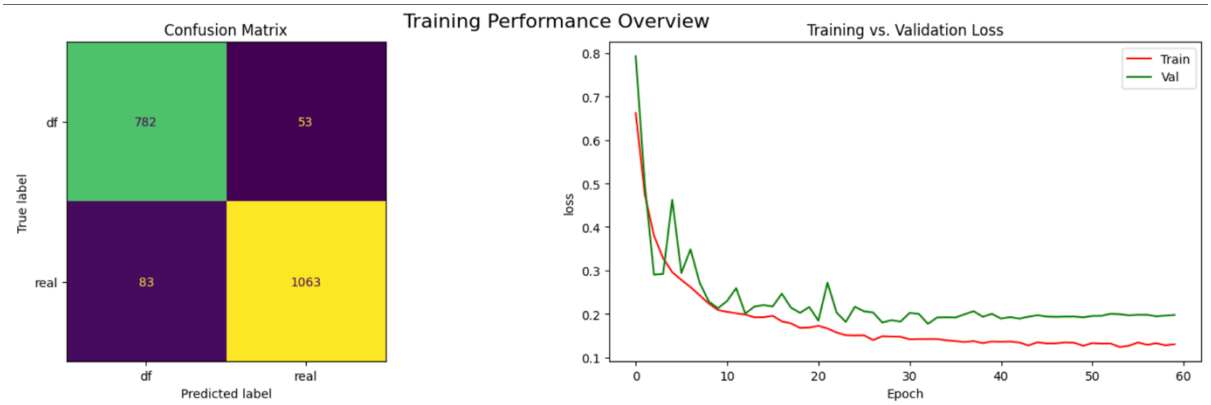
Model Experiments

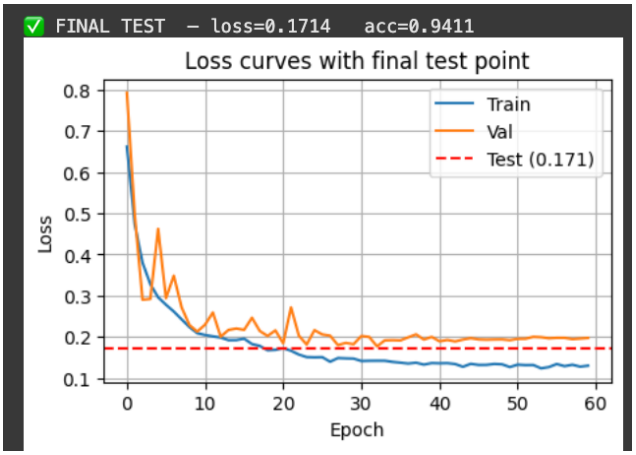




- MesoInception4 Training Results:
  - Best validation accuracy achieved was approximately 93.13% at epoch 28.
  - Final test accuracy was 93.1%.

```
Classification report (threshold = 0.50)
              precision    recall  f1-score   support

          df      0.904     0.937     0.920       835
        real      0.953     0.928     0.940      1146

    accuracy                          0.931      1981
   macro avg      0.928     0.932     0.930      1981
weighted avg      0.932     0.931     0.931      1981
```





These results ensure the robust of the chosen hyperparameter and enhancements, setting a solid foundation for the subsequent benchmark comparison against the original author's model.

Training 3D Convolution based Meso Model

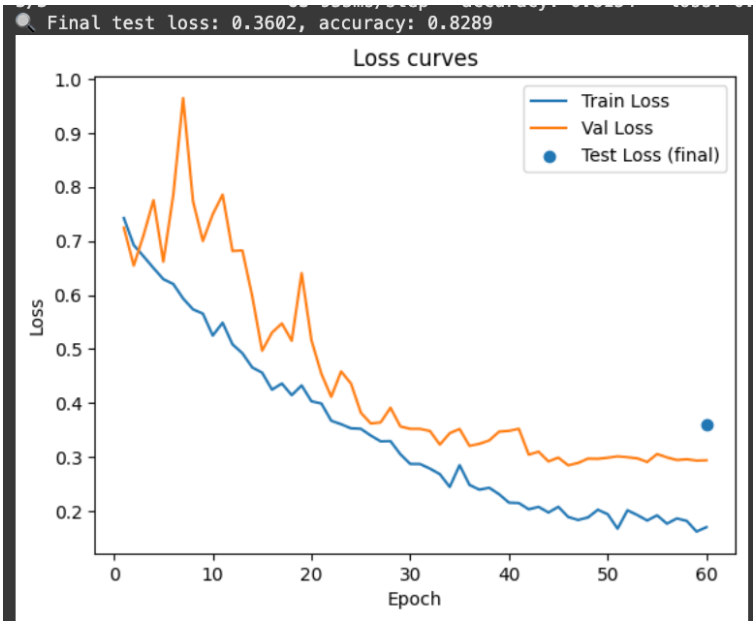
In addition to enhancing and exploring Meso4 and MesoInception4 models, a model was developed which based on 3D convolution instead of 2D convolution and we name it Meso3D model. Meso3D was trained directly on videos (not like Meso4 and MesoInception4 which was trained on cropped frames/ images) with different hyperparameters to identify optimal configurations for improved deepfake detection performance. Multiple training iterations were executed with adjustments in epochs, batch size, and learning rates. Also similar to Meso4 and MesoInception4 models it was trained with AdamW optimizer, and Binary Cross-Entropy loss.

The Integration of 3D Convolution (Meso3D), was aimed to further improve the model's capability to process temporal features in video-based deepfake detection. The initial training parameters for the Meso3D were as follows:

- Meso3D: {Epochs: 60, Batch size: 64, starting learning rate: 0.001, final learning rate reached 0.000063, and Data augmentation: Enabled}

The results from the Meso3D training indicated that best validation accuracy was 86.49% at epoch 36, and test accuracy of 82.89%. Further tuning, computation resources and architectural adjustments could potentially enhance its detection performance.

Overall, visual representations showed the effectiveness of the explored enhancements. These comprehensive experiments establish a solid baseline and provide valuable insights for future model development and optimization.



Benchmarking Against Previous Models

Our enhancements included the implementation of data augmentation techniques, the adoption of Binary Cross-Entropy as the primary loss function, and integrating the AdamW optimizer to achieve improved generalization and performance. In addition to introducing 3D convolution.

1) Using Author's Original Dataset (preprocessing cropped frames .jpg files)

We utilized the authors' original dataset, reorganizing it into new training, validation, and testing subsets to ensure robust evaluation. After training our enhanced model, we conducted a detailed benchmark comparison against the original Meso4 model provided by the authors using identical test conditions and datasets.

Model Experiments

The benchmark results on our test dataset consisting of 1,951 images demonstrated significant performance improvement as shown below:

```
──────── Author weights ────────
Loss = 0.0861 | Acc = 0.8939
Confusion-matrix
 [[ 720   81]
 [ 126 1024]]

Classification-report
              precision    recall  f1-score   support

          df      0.851     0.899     0.874       801
        real      0.927     0.890     0.908      1150

    accuracy                          0.894      1951
   macro avg      0.889     0.895     0.891      1951
weighted avg      0.896     0.894     0.894      1951
```

```
──────── Fine-tuned ────────
Loss = 0.0574 | Acc = 0.9241
Confusion-matrix
 [[ 715   86]
 [  62 1088]]

Classification-report
              precision    recall  f1-score   support

          df      0.920     0.893     0.906       801
        real      0.927     0.946     0.936      1150

    accuracy                          0.924      1951
   macro avg      0.923     0.919     0.921      1951
weighted avg      0.924     0.924     0.924      1951
```

```
========  BENCHMARK SUMMARY  ========
model             |  loss |  acc
------------------------------------
Author weights    | 0.086 | 0.894
Fine-tuned        | 0.057 | 0.924
```

As shown above the classification reports reflected improvements in precision, recall, and F1-score, particularly highlighting the fine-tuned model's ability to distinguish real images from fake ones.

As an enhanced output we have included qualitative evaluations by visually inspecting some samples from the test set as illustrated below. The visual results showed noticeable improvements in prediction confidence and classification correctness when comparing our enhanced fine tuned model to the original Meso4 model. Such visual inspections confirmed the effectiveness of our enhancements, providing clearer differentiation between real and fake images.

Note# 0.5 is the threashold in the classification, where values below 0.5 and near to 0 represent fake class, and values equal or more than 0.5 represent real class.

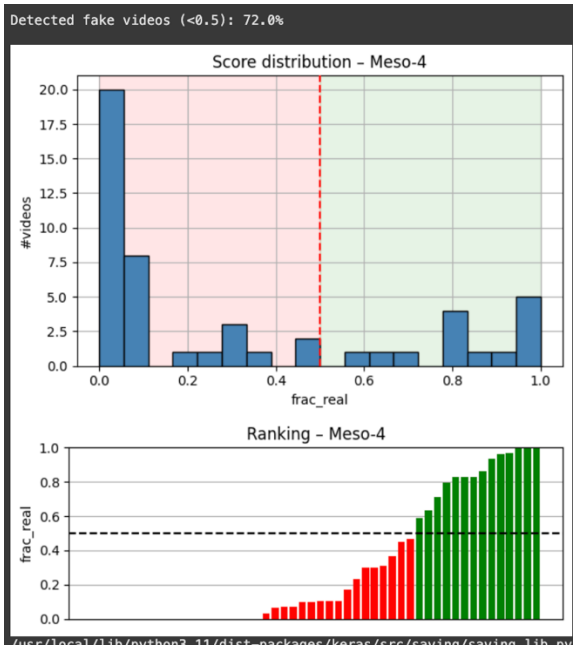
2) Using External FaceForeniscs++ (deepfake videos .mp4 files)

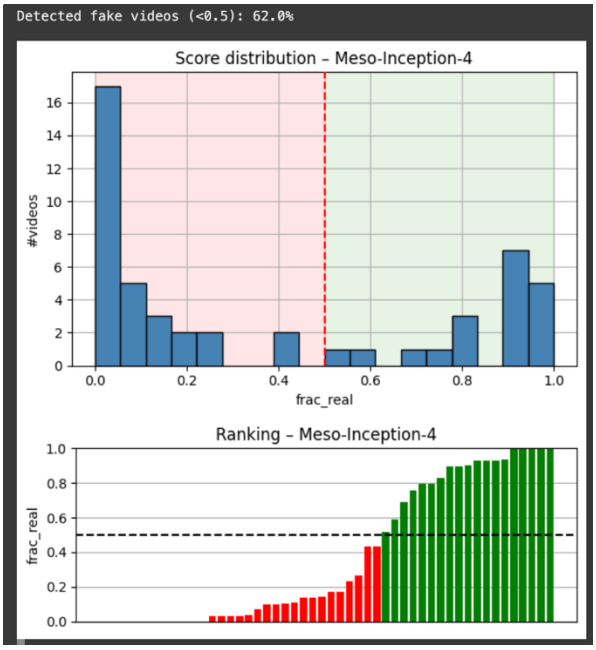
To ensure robust performance we have utilized 50 deepfake videos extracted from FaceForeniscs++ dataset to build our comparision in performance. Below table shows a comparative benchmark against the previously trained versions of Meso4 and MesoInception4 models. In this benchmark both versions were evaluated on using the same test videos:

Dataset	Model	Old Fake Detection Rate	Updated Fake Detection Rate
FaceForensics++ Dataset – 50 Deepfake videos	Meso4	28.0%	72.0%
FaceForensics++ Dataset – 50 Deepfake videos	MesoInception4	12.0%	64.0%

The results show significant improvement in the fake video detection rate, highlighting the improvments in both models.

Model Experiments





Future Research Directions

As we have enjoyed working with the MesoNet for further emplemntation we suggest the following:

1- Enhance the Video and face extraction for videos
2- Enhance Meso3D for the 3D convolutions
3- Train more generalized artifacts such as lip and eyes motion to detect deepfake and other types of faked data for videos and images.
4- Ensure the ensemble learning effectiveness
5- Allow 3D convolutions to handle images too

Conclusion

The comprehensive experimental evaluation provided in this report shows the improvements prospoed from those updated Meso4, MesoInception4 models, and Meso3D models. The benchmarking against the old Meso4 and MesoInception provided by the author, showed that introducing AdamW optimizer, binary cross entroy provided significant enhancement in the generalization to detect new unseen deepfake videos, where with Meso4 the detection rate of unseen data enhanced from 28% to 72% and MesoInception4 from 12% to 64%. These results underscore the efficacy of the refined training methodologies and hyperparameter selections.

The integration of Meso3D model, which involve 3D convolution layers for temporal feature extraction, shows promising results despite slightly lower accuracy, indicating considerable potential for future improvements. These results highlight the value of efficient hyperparameter tuning and advanced training techniques, providing a solid foundation for further advancements in deepfake detection research.

References

[1]] Afchar, D., Nozick, V., Yamagishi, J., & Echizen, I. (2018). Mesonet: A compact facial video forgery detection network. *2018 IEEE International Workshop on Information Forensics and Security (WIFS)*, 1–7. https://doi.org/10.1109/wifs.2018.8630761

[2] Afchar, D., Nozick, V., & Yamagishi, J. (2022). *Evaluation of deepfake detection methods using the Deepfake Detection Challenge dataset*. In *2022 30th European Signal Processing Conference (EUSIPCO)* (pp. 643–647). EURASIP. https://eurasip.org/Proceedings/Eusipco/Eusipco2022/pdfs/0000643.pdf

[3] xdxd003. (n.d.). *FaceForensics++ Dataset (C23)*. Kaggle. https://www.kaggle.com/datasets/xdxd003/ff-c23