



e-ISSN: 2319-8753 | p-ISSN: 2347-6710

IJRSET

International Journal of Innovative Research in
SCIENCE | ENGINEERING | TECHNOLOGY

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

Volume 13, Issue 5, May 2024

ISSN INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA

Impact Factor: 8.423

9940 572 462

6381 907 438

ijirset@gmail.com

www.ijirset.com

Unmasking Deepfake Audio: A Study Using Xception Model

Mugdha Kokate¹, Sandesh Lingayat², Sailaja Rao³, Anish Thube⁴, Shubhangi Motewar⁵

U.G. Student, Dept. of Electronics and Computer Science, Shah & Anchor Kutchhi Engineering College,
Mumbai, India^{1,2,3,4}

Associate Professor, Dept. of Electronics and Computer Science, Shah & Anchor Kutchhi Engineering College,
Mumbai, India⁵

ABSTRACT: Deepfake technology poses a significant threat to digital authenticity, necessitating robust detection mechanisms. The present study addresses the challenge of deepfake audio detection. Utilizing Mel-frequency cepstral coefficients (MFCCs) and the Xception model, a robust detection system has been developed to identify subtle manipulations in audio recordings. Additionally, a user-friendly web application for real-time deepfake audio detection has been deployed. The current focus lies in enhancing the model's performance using advanced techniques. Extensive experimentation and benchmarking against state-of-the-art methods are underway to validate and refine the approach. Anticipated findings will substantially contribute to digital forensics, bolstering defences against audio manipulation in the proliferation of synthetic media.

KEYWORDS: Deepfake, Xception model, MFCC, Deep learning.

I. INTRODUCTION

Deepfakes are like super-powered copycats, using AI to mimic someone's face, voice, or even writing style so convincingly that seem real. Imagine a video where your friend appears to say something they never did or an audio recording where a celebrity's voice is used to promote a fraud, that is the power (and potential danger) of deepfake. Deepfakes are AI-driven impersonators, and are getting more sophisticated all the time. DeepFake poses significant threats across various aspects of life, particularly through their malicious applications. One prominent example is the spread of fake news, where DeepFake technology can be used to create fabricated videos of politicians or public figures making false statements, leading to potential social and political conflicts. Additionally, tampered audio samples can be exploited to deceive speaker recognition systems, such as those used in banking systems for voice-based authentication, thus undermining security measures. Furthermore, DeepFake-based impersonation techniques are also utilized for extortion purposes. The creation of DeepFake audio is relatively straightforward, facilitated by widely available open-source toolkits that require only basic computing resources. Moreover, Real-Time Voice Cloning (RTVC) further simplifies the process, making it accessible to individuals with minimal technical expertise. Audio spoofing is a challenge within speech processing, similar to the concept of DeepFake but with a different target audience. Audio spoofing targets automatic speech verification systems, whereas DeepFake attempts to fool people into thinking audio is real. Unlike DeepFake, which create new content, audio spoofing can manipulate existing samples through methods like replaying or merging audio segments. The growing threat of DeepFake has prompted research into methods for detecting whether an audio recording is genuine or generated. These detection techniques vary in their classification approaches and effectiveness.

There are serious risks to social security, privacy, and authenticity from deepfake. While deepfake films can now be accurately detected thanks to recent advances, specific models are needed to handle fraudulent calls and audio spoofing caused by deepfake. Deepfake audio detection only looks at audio signals, which has gotten less attention than image- and video-based methods [8]. Current methods often leverage both audio and spatiotemporal information from videos, but the classification and detection capabilities of audio-based classifiers are crucial for combating this issue.

The purpose of this work is to devise a technique for separating real audio from deepfake. Enhancing the efficacy of deepfake audio detection, it contributes the following features:

1. Explains how to identify deepfake audio using a method based on transfer learning [10].
2. Expand upon previous research by conducting extensive experiments on the Fake-or-Real dataset and its subsets using both traditional machine learning and deep learning approaches.
3. Extracts the Mel-frequency cepstral coefficients (MFCC) from audio sources using a sophisticated feature extraction technique.

II. RELATED WORK

2.1: Deepfake Audio Detection Techniques Using Deep Convolutional Neural Network

The study "Deepfake Audio Detection Techniques Using Deep Convolutional Neural Network" introduces an innovative tool designed for identifying AI-generated audio deepfakes. This tool leverages a Convolutional Neural Network (CNN) as a robust mechanism for detecting acoustic anomalies, treating the network as a black box. The method is grounded in deep learning principles, presenting a novel approach to audio categorization. Specifically, the paper delves into the utilization of Xception, an architecture derived from the principles of Inception and characterized by Depth Wise Separable Convolutions. Xception aims to optimize the overall system efficiency, thereby enabling real-time detection capabilities. This research positions the proposed models as dependable standard networks within the domain of audio deepfake detection, offering promising avenues for future exploration and application [1].

2.2: SpecRNet: Towards Faster and More Accessible Audio Deepfake Detection

In the paper "SpecRNet: Towards Faster and More Accessible Audio DeepFake Detection," the primary aim is to enhance the accessibility of techniques for detecting audio DeepFake instances [2]. The authors introduce SpecRNet, a neural network architecture distinguished by its efficiency in computational requirements and rapid inference times. Given the proliferation of audio DeepFake utterances generated by deep neural networks, which pose significant risks in various contexts such as the dissemination of fake news, impersonation, and extortion, the development of accessible detection methods is crucial. SpecRNet demonstrates comparable performance to LCNN, a leading model in audio DeepFake detection, while also showcasing up to a 40% reduction in processing time per audio sample, as indicated by the authors' benchmarking. This efficiency makes SpecRNet suitable for adoption by ordinary users on personal devices and by online multimedia platforms for validating a significant volume of daily content.

2.3: Deepfake Audio Detection via MFCC features using Machine Learning

In the paper titled "Deepfake Audio Detection via MFCC Features Using Machine Learning," the focus is on the detection of artificially manipulated or synthesized content, commonly referred to as deepfakes, particularly in the realm of audio. While substantial strides have been made in identifying video-based deepfakes, a notable gap exists in the detection of audio deepfakes. The majority of existing studies have concentrated on this domain, employing various machine learning and deep learning algorithms, often leveraging datasets such as AVSpooof or ASVSpooof.

This study specifically aims to identify deepfake audio content utilizing a combination of machine learning techniques and audio analysis. The core method involves the extraction of pertinent features from the audio signal, particularly leveraging Mel-frequency cepstral coefficients (MFCCs). To ensure the efficacy of the proposed approach, the researchers utilize the Fake-or-Real dataset, which represents the most contemporary benchmark dataset available. This dataset is subdivided into four distinct subsets, namely for-rece, for-2-sec, for-norm, and for-original, with variations in bit rate and audio duration. Notably, these subsets are generated utilizing a text-to-speech model, offering diverse and representative samples for analysis and evaluation.

III. DATASETS UTILIZED

3.1: LJ Speech Dataset

Within the LJSpeech-1.1 dataset, a collection of approximately 13,100 concise audio snippets, totalling around 24 hours of content, is available. These recordings feature a solitary speaker reciting passages extracted from non-fiction texts published between 1884 and 1964. Compiled by the LibriVox project during the years 2016 to 2017, each audio clip is encoded as a single-channel 16-bit PCM WAV file. The segmentation of these clips is automated, primarily based on pauses detected within the recordings. While the boundaries of the clips generally align with the ends of sentences or clauses, this alignment may exhibit variability.

Each audio clip in the dataset is accompanied by a corresponding transcription, meticulously synchronized to ensure accuracy. Furthermore, a quality assurance process has been conducted to validate the alignment of the

transcribed text with the spoken content. It is noteworthy that the original recordings were initially distributed in 128 kbps MP3 format, potentially introducing artefacts associated with MP3 encoding.

These meticulously curated datasets represent invaluable resources for training, validating, and assessing deepfake audio detection models. Offering a diverse array of synthetic and authentic audio samples, they play a pivotal role in the development and evaluation of detection algorithms within the field.

3.2: WaveFake Dataset

The WaveFake dataset is a comprehensive collection of synthetic audio files, all formatted in 16-bit PCM WAV, and comprises around 117,985 artificially generated audio clips. Encompassing approximately 175 hours of diverse audio content, this dataset is readily accessible online for research purposes. It incorporates a wide array of generative architectures, thereby offering a rich spectrum of deepfake audio samples for analysis and experimentation.

Among the notable architectures included in the WaveFake dataset are:

- MelGAN: A feed-forward network operating on Mel spectrograms, featuring a discriminator with multi-resolution capabilities for in-depth feature space analysis.
- ParallelWaveGAN (PWG): Employing a GAN-based training approach with a non-autoregressive version of WaveNet for sample generation, incorporating Short-Time Fourier Transform (STFT) matching for auxiliary loss.
- Multiband MelGAN (MB-MelGAN): Predicts multiple sub-bands, which are then combined to form complete audio, utilizing frequency-based auxiliary loss across various sub-bands.
- Full-band MelGAN (FB-MelGAN): Utilizes L1 distance minimization and feature-space-based loss to directly generate entire audio segments, applying auxiliary loss across the entirety of the audio instead of in sub-bands.
- HiFi-GAN: Minimizes L1 distance between generated and ground truth Mel spectrograms in its loss function, employing multiple sub-discriminators to analyze periodic sections of the input waveform and generate numerous residual blocks simultaneously, thereby capturing patterns of varying lengths.
- WaveGlow: A flow-based generative model inspired by WaveNet, exploring maximum likelihood-based methods [6].

Additionally, the dataset includes samples generated through a complete Text-To-Speech (TTS) pipeline, facilitating the creation of audio based on novel phrases. This wealth of diverse synthetic audio data provided by the WaveFake dataset serves as a valuable resource for the advancement and evaluation of deepfake audio detection techniques.

IV. PREPROCESSING TASKS

4.1: Resampling

In the context of audio deepfake detection, a meticulously curated dataset was compiled by merging the LJ Speech and WaveFake datasets. The primary objective of this endeavour was to establish a comprehensive repository of audio samples tailored for both training and evaluation purposes. Specifically, the LJ Speech snippets were classified as "Bonafide," denoting authentic speech, while the WaveFake samples were categorized as "Fake," signifying manipulated or synthetic content. This amalgamation facilitated the creation of a diverse and representative dataset essential for the development and assessment of deepfake detection algorithms.

4.2: Audio Processing

In the comprehensive examination of audio data, a Voice Activity Detection (VAD) transform played a pivotal role in discerning intricate characteristics. This transformative process meticulously identified and scrutinized segments of silence within the waveform, offering invaluable insights into the structural composition of the audio. Subsequent evaluation of audio quality encompassed considerations of both amplitude characteristics and the durations of silent intervals.

A thorough analysis of the waveform's amplitude provided nuanced insights into the overall robustness and clarity of the audio signal. Consequently, unwanted silences within the audio were judiciously filtered out based on a predefined amplitude threshold, strategically eliminating superfluous pauses and directing focus to segments containing pertinent audio information.

To ensure consistency in processing, all audio samples underwent resampling to a standardized frequency of 16,000 Hz. This standardization facilitated a uniform representation of audio data, thereby streamlining subsequent analysis and model training. The normalization process was meticulously designed to standardize the duration of audio samples, involving either padding or trimming the waveform to achieve a consistent length. This approach enabled effective cross-sample comparison and analysis.

Aligning with predefined temporal specifications, audio waveforms were scrupulously trimmed to match the desired length, ensuring adherence to a consistent temporal standard across all samples in the dataset. This meticulous processing pipeline contributes to the robustness and reliability of subsequent audio deepfake detection models, aligning with rigorous standards in research methodology.

4.3: Visualize Audio

Visual depictions of both the original and processed audio waveforms offered a qualitative insight into the transformations implemented during the pre-processing phases. These visualizations aided in discerning patterns, anomalies, and the overall effects of the pre-processing pipeline.

The following are the waveforms of the original and processed audio:

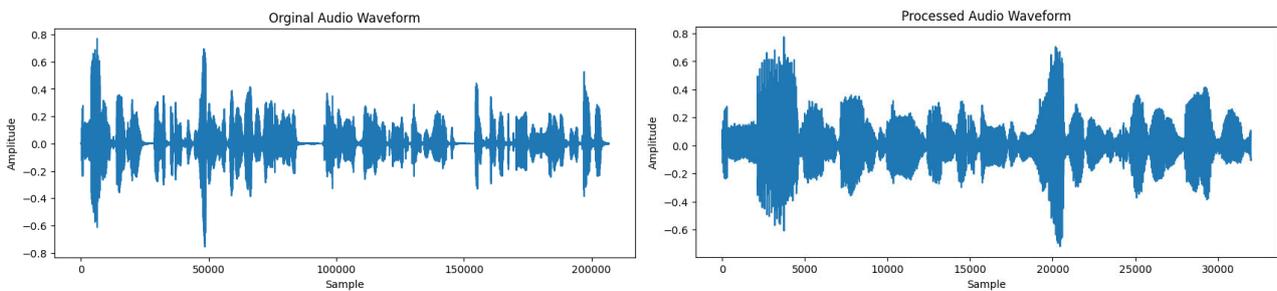


Fig 1.1 Original and Processed Audio Waveforms

The culmination of these extensive pre-processing measures resulted in the refinement and preparation of the dataset for subsequent stages of model development and evaluation in the audio deepfake detection system.

Audio Data Preprocessing

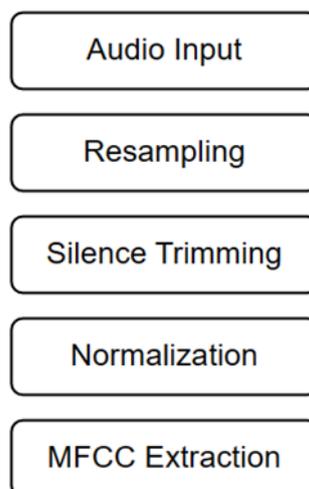


Fig 1.2 Data Pre-processing

V. FEATURE EXTRACTION FOR AUDIO DEEPFAKE DETECTION

In the domain of deepfake audio detection, feature extraction plays a pivotal role, acting as a foundational element in translating raw audio data into meaningful representations for efficient processing by machine learning models. This critical phase significantly impacts the performance of the detection system, necessitating thoughtful selection of the chosen technique to ensure alignment with the objectives and requirements of the model architecture.

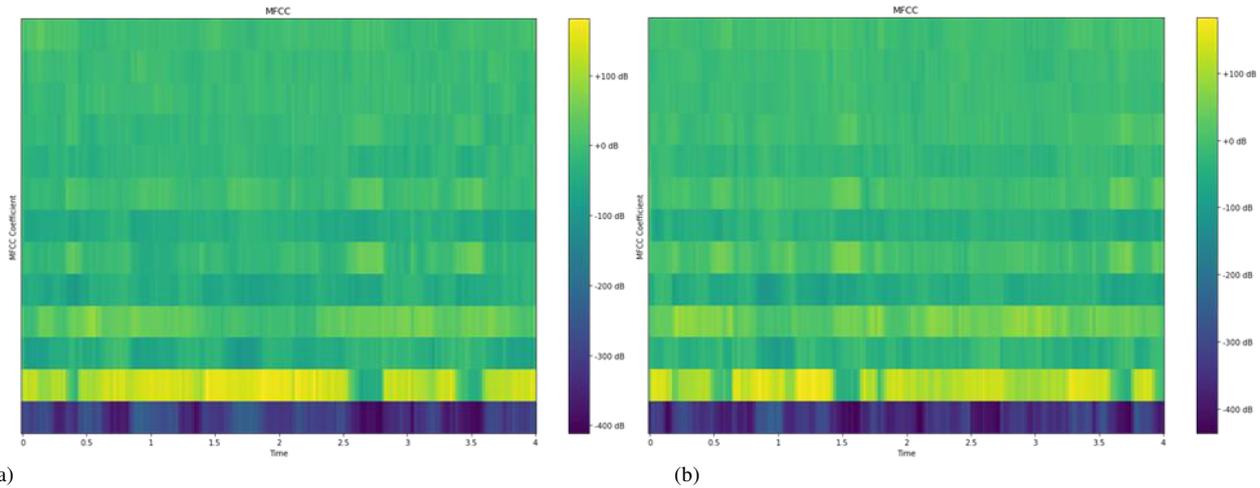


Fig 1.3 (a) MFCC Heatmaps of Processed Bonafide Audio Files
(b) MFCC Heatmaps of Processed Fake Audio Files

1) Mel-Frequency Cepstral Coefficients (MFCC) Extraction

Mel-Frequency Cepstral Coefficients (MFCCs) have emerged as a prevalent technique in audio signal processing, recognized for their effectiveness in capturing spectral characteristics, especially in tasks related to speech. MFCCs provide a succinct yet informative representation of the spectral envelope of an audio signal, making them well-suited for various applications such as speech recognition and audio classification [5].

2) Extraction Process

The extraction procedure involves utilizing the Librosa library to compute Mel-Frequency Cepstral Coefficients (MFCCs) from audio files, achieving a delicate balance between feature richness and computational efficiency. The fine-tuning of parameters such as the number of MFCCs, Fast Fourier Transform (FFT) size, and hop length is essential to optimize performance. Following this, mean aggregation is applied to the computed MFCCs to succinctly capture the essence of each audio file, efficiently condensing temporal information into a representative feature vector. This approach enhances the system's efficacy in processing image-based inputs for precise deepfake audio detection.

This conversion process enriches the representation of audio features, thereby facilitating improved compatibility and performance of the deepfake audio detection system with the Xception model architecture.

VI. XCEPTION MODEL

The Xception model, an abbreviation for "Extreme Inception", stands out as a deep neural network architecture renowned for its effectiveness in various computer vision tasks, including image subsequent analyses within the deepfake audio detection pipeline.

This meticulous feature extraction methodology lays the groundwork for robust deepfake audio detection, empowering the system to discern subtle differences and accurately identify manipulated audio content.

6.1: Conversion to MFCC Heatmaps for Xception Model Compatibility

The Xception model, which has been selected for deepfake audio detection research, operates on image-based inputs. To overcome this constraint and ensure compatibility with the Xception model, the conversion of MFCC features into MFCC heatmaps is proposed [7]. This conversion involves visualizing the MFCC coefficients as heatmaps, where each heatmap represents the intensity of MFCC coefficients across time and frequency bins. By converting MFCC features into heatmaps, the audio data is transformed into a format that aligns with the input requirements of the Xception model. This adaptation enables seamless integration of audio data into the Xception architecture, allowing us to leverage recognition and feature extraction. Originating from the Inception architecture, Xception introduces a novel concept by replacing standard convolutional layers with depth-wise separable convolutions, aiming to capture intricate patterns and dependencies within the data.

6.2: Objective

The objective is to utilize the Xception model to create an effective deep-learning architecture tailored for detecting deepfake audio. This involves leveraging MFCC heatmap images, meticulously processed representations of audio data, during training. The aim is to enhance the model's ability to identify subtle audio features, thus improving its effectiveness in detecting and addressing potential threats posed by deepfake audio instances.

6.3: Data Loading and Preprocessing

Essential libraries such as os, numpy, tensorflow, keras, sklearn, and pandas are utilized. MFCC Heatmaps are promptly converted into arrays and paired with corresponding labels ('Bonafide' or 'fake'). Standardization is achieved by resizing the images to (299, 299) pixels. The dataset is then split into training, validation, and test sets using the `train_test_split` function from sklearn, ensuring meticulous data preparation for further training and validation.

6.4: Model Architecture

The Xception model, pre-trained on ImageNet, is adapted for audio analysis. A custom model is created by removing the top layers and adding classification layers. The architecture incorporates global average pooling, a dense layer with 1024 units, and a final dense layer for binary classification ('Bonafide' or 'Fake'). The model is compiled using the Adam optimizer and sparse categorical cross-entropy loss function.

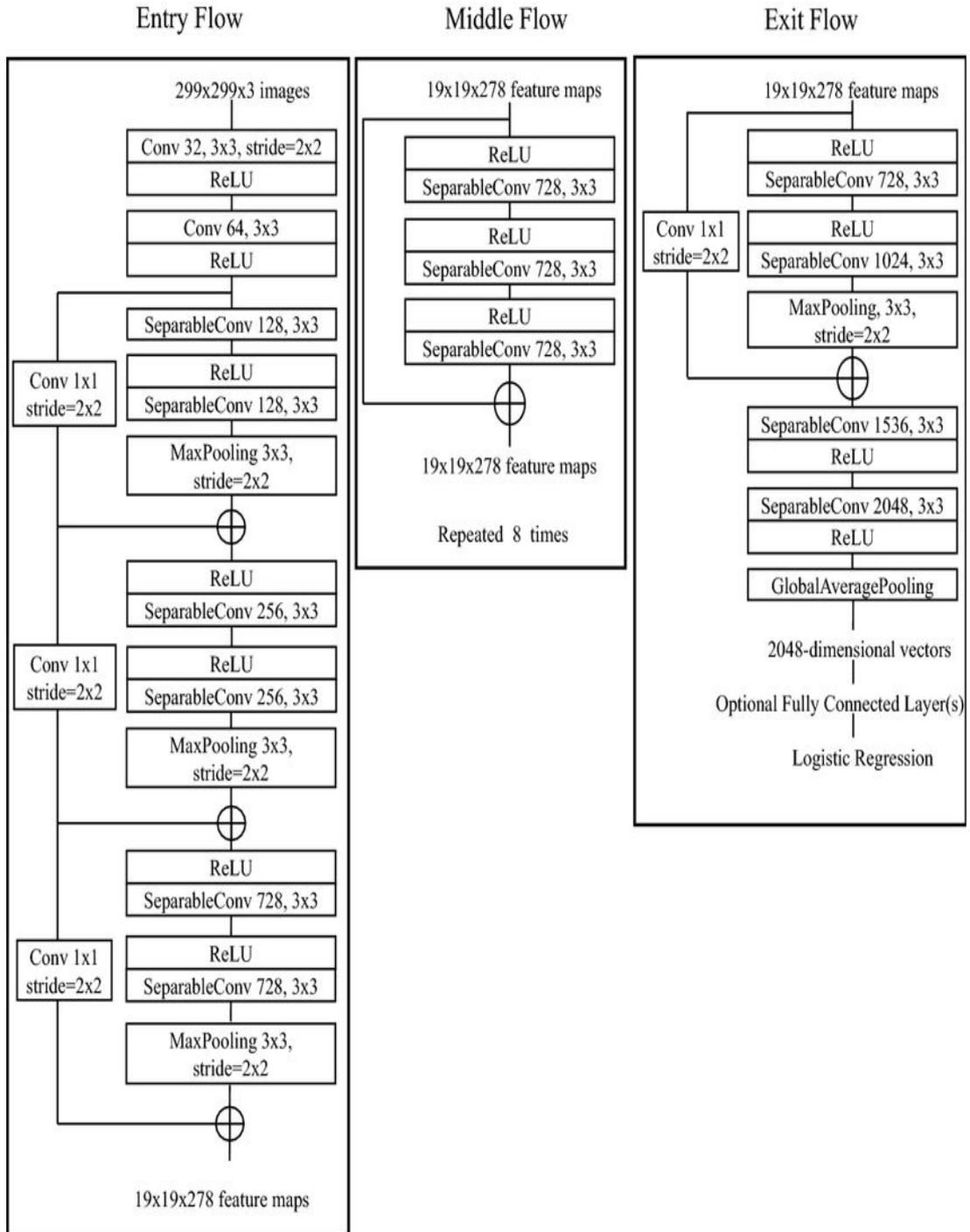


Fig 1.4 Xception architecture [Source: [11]]

6.5: Flowchart

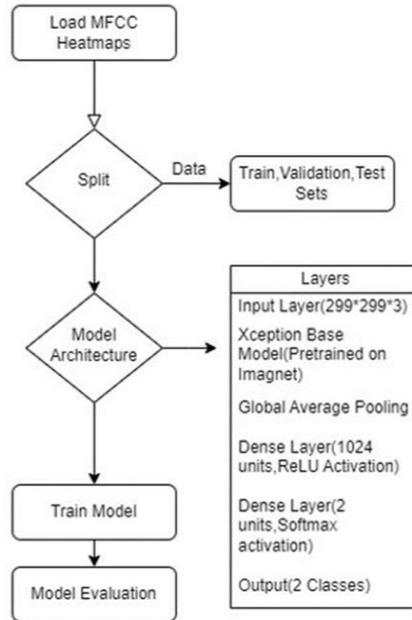


Fig 1.5 Xception Model Training

6.6: Training the Model

During training, the model efficiently incorporates callbacks to optimize the training process. Key callbacks such as Model Checkpoint, Early Stopping, and Reduce LR On Plateau are integrated to ensure model generalization. These callbacks monitor validation accuracy, save the best model weights, prevent overfitting, and dynamically adjust the learning rate if validation accuracy plateaus. Progress, loss, and accuracy metrics are displayed during each epoch. The training history is recorded in a structured manner, promoting optimal convergence and preventing overfitting.

These strategies ensure the model's robustness and its ability to generalize well to unseen data for reliable deepfake audio detection.

VII. MODEL EVALUATION

The model undergoes rigorous evaluation on the test set, providing a comprehensive assessment of its performance metrics.

Evaluation Results:

- Test Accuracy: Achieving a notable accuracy of 70.56%, the model effectively distinguishes between bonafide and deepfake audio samples, demonstrating its robust discriminatory capability.
 - Test Loss: The computed test loss, standing at 0.6129, represents the deviation between the model's predictions and the actual values in the test dataset. This metric provides insights into the model's predictive accuracy.
- The evaluation function not only quantifies the model's accuracy but also generates a detailed classification report and a confusion matrix. These additional analyses offer a nuanced understanding of the model's performance, highlighting its strengths and potential areas for improvement. This set of evaluation metrics collectively affirms the model's competence in deepfake audio detection, providing valuable insights for its deployment in real-world scenarios.

Model Training and Evaluation Results

The deepfake audio detection model was trained and evaluated using a convolutional neural network architecture. The model was trained over 10 epochs, utilizing an initial learning rate 0.0001.



Training Metrics

The training process exhibited the following trends:

	loss	accuracy	val_loss	val_accuracy	lr
0	3.812656	0.535446	1.213851	0.580381	0.0001
1	1.314354	0.570211	0.731165	0.638965	0.0001
2	1.486739	0.564758	2.438244	0.510218	0.0001
3	1.413267	0.578732	0.811995	0.643052	0.0001
4	1.014353	0.600545	1.487107	0.542916	0.0001
5	1.448530	0.594240	0.826251	0.680518	0.0001
6	1.555830	0.585037	1.404120	0.587193	0.0001
7	0.977747	0.633265	0.622360	0.684605	0.0001
8	1.003376	0.627301	1.187919	0.583106	0.0001
9	1.056413	0.617417	0.611001	0.708447	0.0001

Fig 1.6 Training Metrics

Model Evaluation

The evaluation of the deepfake audio detection model on a test dataset resulted in an accuracy of 70.56% and a test loss of 0.6129, demonstrating its ability to discern bonafide and fake audio samples. Insights from the classification report further illustrate the model's performance, indicating its effectiveness in distinguishing between bonafide and fake audio, with a balanced identification of fake samples.

Classification Report:

	precision	recall	f1-score	support
0	0.75	0.61	0.67	914
1	0.67	0.80	0.73	920
accuracy			0.71	1834
macro avg	0.71	0.71	0.70	1834
weighted avg	0.71	0.71	0.70	1834

Fig 1.7 Classification Report

Precision-Recall Curve Analysis

Precision-recall curves are employed to evaluate the model's performance, especially for imbalanced datasets. The trade-off between recall and precision for different thresholds is shown by the curve.

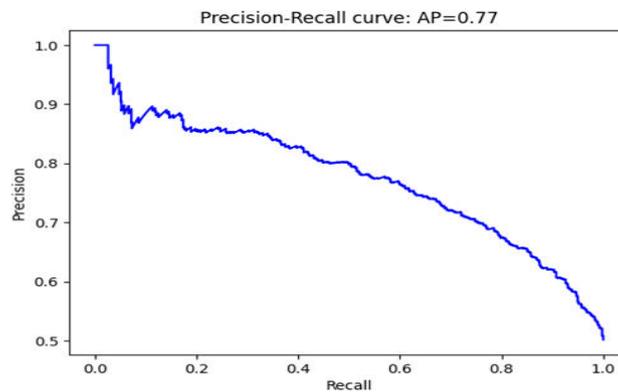


Fig 1.8 Precision-Recall Curve

Confusion Matrix Analysis

The confusion matrix illustrates the model's ability to distinguish between bonafide and fake audio samples. It showcases counts of true positives, true negatives, false positives, and false negatives. For instance, the model correctly identified 555 bonafide audio samples but misclassified 359 bonafide samples as fake. Similarly, it correctly predicted 739 fake audio samples and misclassified 181 fake samples as bonafide.

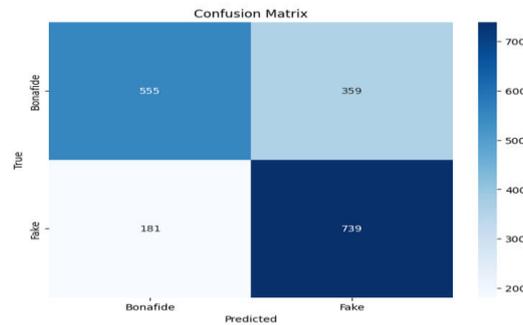


Fig 1.9 Confusion Matrix

VIII. CONCLUSION

The Xception model, coupled with MFCC heatmaps, proves to be a powerful tool for deepfake audio detection. With a test accuracy of 70.56%, the model displays promising results, demonstrating its ability to distinguish between genuine and deepfake audio. The precision-recall curve and confusion matrix analyses provide valuable insights into the model's strengths and areas for potential improvement.

Future Directions

Future work may involve refining the model architecture, exploring additional feature representations, and expanding the dataset. Further enhancements could contribute to even greater accuracy and robustness in the detection of evolving deepfake audio techniques.

IX. FUTURE SCOPE

The future of deepfake audio detection research holds promising avenues for further exploration and enhancement. Building upon the current study, several potential directions are already underway:

1. **Real-time Detection Implementation:** Development of real-time deepfake audio detection systems is in progress, aiming to address the rapid dissemination of manipulated content across digital platforms. Integration of advanced algorithms and optimization techniques is actively being pursued to enhance processing speed without compromising accuracy.
2. **Enhanced Model Architectures:** Ongoing research involves exploring novel model architectures and leveraging advancements in deep learning, such as attention mechanisms and transformer-based architectures, to improve detection accuracy and robustness. Hybrid models combining multiple architectures are currently being investigated to harness synergistic benefits.
3. **Adversarial Training Strategies:** Work has begun on implementing adversarial training methods to fortify models against sophisticated attacks. Adversarial robustness training aims to enhance the resilience of deepfake detection models by exposing them to diverse manipulation techniques during training.
4. **Development of Android Application:** As part of ongoing efforts, an Android application is actively being developed for deepfake audio detection. This application aims to provide users with a user-friendly interface to detect and verify the authenticity of audio content on mobile devices.
5. By advancing these research endeavors, the goal is to propel the field of deepfake audio detection forward, fostering a safer and more trustworthy digital ecosystem for all individuals.

REFERENCES

1. B. Kumar and S. R. Alraisi, "Deepfake Audio Detection Techniques Using Deep Convolutional Neural Network," 2022 International Conference on Machine Learning, Big Data, Cloud and Parallel Computing (COM-IT-CON), Faridabad, India, 2022, pp. 463-468, doi: 10.1109/COM-IT-CON54601.2022.9850771.
2. P. Kawa, M. Plata and P. Syga, "SpecRNet: Towards Faster and More Accessible Audio DeepFake Detection," 2022 IEEE International Conference on Trust, Security and Privacy in Computing and Communications (TrustCom), Wuhan, China, 2022, pp. 792-799, doi: 10.1109/TrustCom56396.2022.00111.
3. Hamza et al., "Deepfake Audio Detection via MFCC Features Using Machine Learning," in IEEE Access, vol. 10, pp. 134018-134028, 2022, doi: 10.1109/ACCESS.2022.3231480.

4. Yi, J., Bai, Y., Tao, J., Ma, H., Tian, Z., Wang, C., Wang, T., & Fu, R. (2023). Half-Truth: A Partially Fake Audio Detection Dataset. arXiv:2104.03617 [cs.SD]
5. Khochare, J., Joshi, C., Yenarkar, B. et al. A Deep Learning Framework for Audio Deepfake Detection. Arab J Sci Eng 47, 3447–3458 (2022). <https://doi.org/10.1007/s13369-021-06297-w>
6. Almutairi, Z.; Elgibreen, H. A Review of Modern Audio Deepfake Detection Methods: Challenges and Future Directions. Algorithms 2022, 15, 155. <https://doi.org/10.3390/a15050155>
7. K. Li, X. Lu, M. Akagi and M. Unoki, "Contributions of Jitter and Shimmer in the Voice for Fake Audio Detection," in IEEE Access, vol. 11, pp. 84689-84698, 2023, doi: 10.1109/ACCESS.2023.3301616.
8. Z. M. Almutairi and H. Elgibreen, "Detecting Fake Audio of Arabic Speakers Using Self-Supervised Deep Learning," in IEEE Access, vol. 11, pp. 72134-72147, 2023, doi: 10.1109/ACCESS.2023.3286864.
9. R. L. M. A. P. C. Wijethunga, D. M. K. Matheesha, A. A. Noman, K. H. V. T. A. De Silva, M. Tissera and L. Rupasinghe, "Deepfake Audio Detection: A Deep Learning Based Solution for Group Conversations," 2020 2nd International Conference on Advancements in Computing (ICAC), Malabe, Sri Lanka, 2020, pp. 192-197, doi: 10.1109/ICAC51239.2020.9357161.
10. Camacho, S., Ballesteros, D.M., Renza, D. (2021). Fake Speech Recognition Using Deep Learning. In: Figueroa-García, J.C., Díaz-Gutierrez, Y., Gaona-García, E.E., Orjuela-Cañón, A.D. (eds) Applied Computer Sciences in Engineering. WEA 2021. Communications in Computer and Information Science, vol 1431. Springer, Cham. https://doi.org/10.1007/978-3-030-86702-7_4
11. https://www.researchgate.net/figure/Xception-architecture-used-as-CNN-module-in-the-proposed-work_fig16_349310566



INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

 9940 572 462  6381 907 438  ijirset@gmail.com



www.ijirset.com

Scan to save the contact details