# MFCCs, Chroma Features, and Spectrogram Images for Deepfake Audio Classification Using Machine Learning

Nora Bakken, Shavil Singh, Makwana Prashant Lnu, and Tapadhir Das Department of Computer Science, University of the Pacific Stockton, California, USA n bakken@u.pacific.edu,tdas@pacific.edu

#### **ABSTRACT**

#### **CCS CONCEPTS**

• Mel-frequency cepstral coefficients (MFCCs) → Audio representation; • Convolutional Neural Networks → Machine Learning; • Adversarial attack protection → Cybersecurity.

### **KEYWORDS**

MFCCs, Spectrogram, VGG16, ResNet50, Chroma Features, SVM, Gradient Boosting, Deepfakes

#### **ACM Reference Format:**

#### 1 INTRODUCTION

#### 2 MOTIVATION

With the rise of artificial intelligence (AI), deepfakes are more prevalent than ever, bringing with them a slough of potential dangers in a variety of areas. Likely the most well-known effect of deepfakes in everyday life is the rise of false media, especially targeting individuals.

The American Bar Association highlighted a targeted defamation attack involving an audio recording with the voice of a high school principal making racist and antisemitic comments. After the recording spread throughout the school, the principal was in danger of losing his livelihood. He denied making these comments. After a thorough investigation, the local police deemed the recording to have been manipulated using AI [Jr. 2024].

The less talked about consequence of accessible and easy to create deepfakes is the rise of non-consensual explicit deepfake attacks on individuals. These attacks, along with being traumatic to the individual, are expanding the ever-present gender gap at the global level, inflicting consequences at the societal level [Kim 2024].

In the recent United States election, nearly half of American voters stated that deepfakes had an influence on their ballots [Genovese

## Unpublished working draft. Not for distribution.

for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, o republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

Conf 2024, Date, Place

© 2024 Copyright held by the owner/author(s). Publication rights licensed to ACM. ACM ISBN XXX-X-XXXXX-XXXXX-X/XX/XX

loss due to a deepfake [Zaki 2024]. This is a significant economic impact.

These are just a few of many examples of individuals who have been hurt by deepfake attacks. There is no question about the

2024]. Looking at this problem from a purely monetary perspective,

CFO states that 92 percent of companies have experienced financial

These are just a few of many examples of individuals who have been hurt by deepfake attacks. There is no question about the negative impact that deepfakes actively have on our lives. It is now imperative that an application is developed to reliably identify deepfakes so that fewer individuals are harmed.

#### 3 LITERATURE REVIEW

#### 4 SYSTEM MODEL / BACKGROUND

Numerous technologies and algorithms were explored to gather insights on the most effective methods for deepfake audio detection. In this section we will explore the complex features we utilized in our data preprocessing pipeline and discuss the sophisticated machine learning models we employed in our experiments.

## 4.1 MFCCs

Mel-Frequency Cepstral Coefficients, or MFCCs, are a widely used feature set for speech recognition and other audio processing applications. They are a competitive feature set because they mimic what the human ear perceives [Hamza et al. 2022].

"MFCC coefficients are obtained by de-correlating the output log energies of a filter bank which consists of triangular filters, linearly spaced on the Mel frequency scale. Conventionally an implementation of discrete cosine transform (DCT) known as distributed DCT (DCT - II) is used to de-correlate the speech as it is the best available approximation of the Karhunen-Loeve Transform (KLT)" [Hossan et al. 2010].

The resulting vector represents spectral acoustic features as floating-point numbers. These values capture essential characteristics of the audio that aid the model in distinguishing between real and fake samples. The vector's size is flexible and depends on user configuration.

Generating a vector of MFCC features was an essential step in our preprocessing efforts. In our experiments, we used a sampling rate of 22050 Hz, which resulted in 20 features for each audio sample. These features were then saved to a CSV file, making them available for later model training. These features played a pivotal role in the model's ability to make accurate decisions.

#### 4.2 Mel-Spectrogram Images

Nora TODO

2024-12-04 12:52. Page 1 of 1-2.

Ameer Hamza, Abdul Rehman Rehman Javed, Farkhund Iqbal, Natalia Kryvi Ahmad S. Almadhor, Zunera Jalil, and Rouba Borghol. 2022. Deepfake A Detection via MFCC Features Using Machine Learning. IEEE Access 10 (2 134018–134028. https://doi.org/10.1109/ACCESS.2022.3231480  Md. Afzal Hossan, Sheeraz Memon, and Mark A Gregory. 2010. A novel approac MFCC feature extraction. In 2010 4th International Conference on Signal Proce and Communication Systems. 1–5. https://doi.org/10.1109/ICSPCS.2010.5709  METHODOLOGY  RESULTS  FUTURE WORK  Jora TODO  Same and Convertible of explicit deepfakes wrecks womens and deepens gender divide. https://www.pbs.org/newshour/world/in-south-k rise-of-explicit-deepfakes-wrecks-womens-lives-and-deepens-gender-divide cessed: 2024-12-03. Adam Zaki. 2024. 92to a deepfake. https://www.cfo.com/news/most-compa have-experienced-financial-loss-due-to-a-deepfake-regula-report/732094/?	Conference Data Black	
Ameer Hamza, Abdul Rehman Rehman Javed, Farkhund Iqbal, Natalia Kryvi Ahmad S. Almadhor, Zunera Jalil, and Rouba Borghol. 2022. Deepfake A Detection via MFCC Features Using Machine Learning. IEEE Access 10 (2 134018–134028. https://doi.org/10.1109/ACCESS.2022.3231480   Md. Afzal Hossan, Sheeraz Memon, and Mark A Gregory. 2010. A novel approac and Communication Systems. 1–5. https://doi.org/10.1109/ICSPCS.2010.5709   Judge Herbert B. Dixon Jr. 2024. The "Deepfake Defense": An Evidentiary Codrum. https://www.americanbar.org/groups/judicial/publications/judges_jou 2024/spring/deepfake-defense-evidentiary-conundrum/. Accessed: 2024-12-03.  RESULTS   FUTURE WORK   Jora TODO   Adam Zaki. 2024. 1. South Korea, rise of explicit deepfakes wrecks women's and deepens gender divide. https://www.pbs.org/newshour/world/in-south-k rise-of-explicit-deepfakes-wrecks-womens-lives-and-deepens-gender-divide cessed: 2024-12-03.  Adam Zaki. 2024. 92to a deepfake. https://www.cfo.com/news/most-compa have-experienced-financial-loss-due-to-a-deepfake-regula-report/732094/? campaign=Yahoo-Licensed-Content&utm_source=yahoo&utm_medium=ref Accessed: 2024-12-03.  REFERENCES   Received XX December 2024; revised XX XXXX; accepted X X XXXX accepted X X XXXX.	Conf 2024, Date, Place	Bakken, Singh, & Lnu, et
Ahmad S. Almadhor, Zunera Jalil, and Rouba Borghol. 2022. Deepfake A Detection via MFCC Features Using Machine Learning. IEEE Access 10 (2) 134018–134028. https://doi.org/10.1109/ACCESS.2022.3231480 Md. Afzal Hossan, Sheeraz Memon, and Mark A Gregory. 2010. A novel approad MFCC feature extraction. In 2010 4th International Conference on Signal Proce and Communication Systems. 1–5. https://doi.org/10.1109/ICSPCS.2010.5709 Judge Herbert B. Dixon Jr. 2024. The "Deepfake Defense": An Evidentiary Confurm. https://www.americanbar.org/groups/judicial/publications/judges. 2024/spring/deepfake-defense-evidentiary-conundrum/. Accessed: 2024-12-Wign-Jin Kim. 2024. In South Korea, rise of explicit deepfakes wrecks women's and deepens gender divide. https://www.pbs.org/newshour/world/in-south-krise-of-explicit-deepfakes-wrecks-womens-lives-and-deepens-gender-divide cessed: 2024-12-03.  ACKNOWLEDGMENTS  REFERENCES  Aniella Genovese. 2024. Nearly 50% of voters said deepfakes had some influence on	4.3 Convolutional Neural Networks	said-deepfakes-had-some-influence-election-decision. Accessed: 2024-12-03.
134018–134028. https://doi.org/10.1109/ACCESS.2022.3231480  Md. Afzal Hossan, Sheeraz Memon, and Mark A Gregory. 2010. A novel approac MFCC feature extraction. In 2010 4th International Conference on Signal Proce and Communication Systems. 1–5. https://doi.org/10.1109/ICSPCS.2010.5709  METHODOLOGY  MADA INTERNATION AND METHODOLOGY  MICHODOLOGY  MICHODOLOGY  MICHODOLOGY  MICHODOLOGY  MICHODOLO	Jora TODO	Ahmad S. Almadhor, Zunera Jalil, and Rouba Borghol. 2022. Deepfake Au
MFCC feature extraction. In 2010 4th International Conference on Signal Proce and Communication Systems. 1–5. https://doi.org/10.1109/ICSPCS.2010.5709  METHODOLOGY  METHODOLOGY  METHODOLOGY  METHODOLOGY  METHODOLOGY  METHODOLOGY  Judge Herbert B. Dixon Jr. 2024. The "Deepfake Defense": An Evidentiary Condrum. https://www.americanbar.org/groups/judicial/publications/judges_jou 2024/spring/deepfake-defense-evidentiary-conundrum/. Accessed: 2024-12-Hyung-Jin Kim. 2024. In South Korea, rise of explicit deepfakes wrecks women's and deepens gender divide. https://www.pbs.org/newshour/world/in-south-krise-of-explicit-deepfakes-wrecks-womens-lives-and-deepens-gender-divide cessed: 2024-12-03.  Adam Zaki. 2024. 92to a deepfake. https://www.cfo.com/news/most-compa have-experienced-financial-loss-due-to-a-deepfake-regula-report/732094/? campaign=Yahoo-Licensed-Content&utm_source=yahoo&utm_medium=ref Accessed: 2024-12-03.  REFERENCES  Paniella Genovese. 2024. Nearly 50% of voters said deepfakes had some influence on	4.3.1 VGG16. Nora TODO	<ul> <li>Md. Afzal Hossan, Sheeraz Memon, and Mark A Gregory. 2010. A novel approach MFCC feature extraction. In 2010 4th International Conference on Signal Process. and Communication Systems. 1–5. https://doi.org/10.1109/ICSPCS.2010.57097.</li> <li>Judge Herbert B. Dixon Jr. 2024. The "Deepfake Defense": An Evidentiary Condrum. https://www.americanbar.org/groups/judicial/publications/judges_jour. 2024/spring/deepfake-defense-evidentiary-conundrum/. Accessed: 2024-12-01.</li> <li>Hyung-Jin Kim. 2024. In South Korea, rise of explicit deepfakes wrecks women's liand deepens gender divide. https://www.pbs.org/newshour/world/in-south-korise-of-explicit-deepfakes-wrecks-womens-lives-and-deepens-gender-divide. ccssed: 2024-12-03.</li> <li>Adam Zaki. 2024. 92to a deepfake. https://www.cfo.com/news/most-companhave-experienced-financial-loss-due-to-a-deepfake-regula-report/732094/?ut campaign=Yahoo-Licensed-Content&amp;utm_source=yahoo&amp;utm_medium=reference</li> </ul>
Judge Herbert B. Dixon Jr. 2024. The "Deepfake Defense": An Evidentiary Condrum. https://www.americanbar.org/groups/judicial/publications/judges_jou 2024/spring/deepfake-defense-evidentiary-conundrum/. Accessed: 2024-12-Hyung-Jin Kim. 2024. In South Korea, rise of explicit deepfakes wrecks women's and deepens gender divide. https://www.pbs.org/newshour/world/in-south-k rise-of-explicit-deepfakes-wrecks-womens-lives-and-deepens-gender-divide cessed: 2024-12-03.  Adam Zaki. 2024. 92to a deepfake. https://www.cfo.com/news/most-compa have-experienced-financial-loss-due-to-a-deepfake-regula-report/732094/? campaign=Yahoo-Licensed-Content&utm_source=yahoo&utm_medium=ref Accessed: 2024-12-03.  REFERENCES  Paniella Genovese. 2024. Nearly 50% of voters said deepfakes had some influence on	4.3.2 ResNet50. Nora TODO	
drum. https://www.americanbar.org/groups/judicial/publications/judges_jou 2024/spring/deepfake-defense-evidentiary-conundrum/. Accessed: 2024-12-8	5 METHODOLOGY	
FUTURE WORK NOTA TODO  FUTURE WORK NOTA TODO  S CONCLUSION ACKNOWLEDGMENTS  REFERENCES  Paniella Genovese. 2024. Nearly 50% of voters said deepfakes had some influence on  Hyung-Jin Kim. 2024. In South Korea, rise of explicit deepfakes wrecks women's and deepens gender divide. https://www.pbs.org/newshour/world/in-south-k rise-of-explicit-deepfakes-wrecks-womens-lives-and-deepens-gender-divide cessed: 2024-12-03.  Adam Zaki. 2024. 92to a deepfake. https://www.cfo.com/news/most-compa have-experienced-financial-loss-due-to-a-deepfake-regula-report/732094/? campaign=Yahoo-Licensed-Content&utm_source=yahoo&utm_medium=ref Accessed: 2024-12-03.  Received XX December 2024; revised XX X XXXX; accepted X X XXXX		
Fise-of-explicit-deepfakes-wrecks-womens-lives-and-deepens-gender-divide cessed: 2024-12-03.  Adam Zaki. 2024. 92to a deepfake. https://www.cfo.com/news/most-compa have-experienced-financial-loss-due-to-a-deepfake-regula-report/732094/? campaign=Yahoo-Licensed-Content&utm_source=yahoo&utm_medium=ref Accessed: 2024-12-03.  REFERENCES  Paniella Genovese. 2024. Nearly 50% of voters said deepfakes had some influence on		
Adam Zaki. 2024. 92to a deepfake. https://www.cfo.com/news/most-compa have-experienced-financial-loss-due-to-a-deepfake-regula-report/732094/? campaign=Yahoo-Licensed-Content&utm_source=yahoo&utm_medium=ref Accessed: 2024-12-03.  REFERENCES  Paniella Genovese. 2024. Nearly 50% of voters said deepfakes had some influence on		
CONCLUSION  campaign=Yahoo-Licensed-Content&utm_source=yahoo&utm_medium=ref ACKNOWLEDGMENTS  REFERENCES  Paniella Genovese. 2024. Nearly 50% of voters said deepfakes had some influence on		
REFERENCES  Received XX December 2024; revised XX X XXXX; accepted X X XXXX  Paniella Genovese. 2024. Nearly 50% of voters said deepfakes had some influence on		
Daniella Genovese. 2024. Nearly 50% of voters said deepfakes had some influence on		CX.
		Received AX December 2024; revised AX A AAAA; accepted A A AAA
Jun Holistical West High		difficultion.
JRIP HOLE OF HISHIP	1	
The distribution of the di		
Jrip Hot Gir		XX
JAP HOLEST CO.		
THE HOLE OF		
and the following the second s	113	
Jriph-Jot ,	10 x 80,	